

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2002, 20:55:58 ; Search time 1486.74 Seconds  
(without alignments)  
8651.612 Million cell updates/sec

Title: US-09-842-316-1

Perfect score: 1197

Sequence: 1 atggagtcgggctgctgcg.....cagaaccggctgcagactga 1197

Scoring table:

IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_esthum:\*  
4: em\_esthum:\*  
5: em\_esthum:\*  
6: em\_esthum:\*  
7: em\_esthum:\*  
8: em\_esthum:\*  
9: em\_esthum:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_est3:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pin:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	568.8	47.5	757	11	BI181178
C 2	346.2	28.9	487	10	AW488801
C 3	326	27.2	725	11	BF383651
C 4	292.8	24.5	763	11	BI329992
C 5	291.8	24.4	295	11	BF333652
C 6	271.2	22.7	3015	12	AK004591
C 7	252	21.1	973	13	CNS02502
C 8	242.2	20.2	1126	11	BF383651
C 9	223.6	18.7	960	13	CNS01WAS
C 10	216.6	18.1	694	11	BI332488
C 11	216.6	18.1	793	11	BF383651
C 12	216.6	18.1	886	10	AL560216

13	215.2	18.0	907	10	AL560608
14	208	17.4	745	11	BF281109
15	197.2	16.5	695	11	BF281107
16	196.2	15.9	924	13	CNS02REF
17	175.4	15.0	347	10	AI317881
18	174.4	14.6	281	10	BB567592
19	174.4	14.6	460	10	BE55413
20	173.2	14.5	912	11	BE540337
21	168	14.0	535	11	BF188941
22	166.2	13.9	779	11	BF383651
23	155.6	13.3	707	10	AL543816
24	156.8	13.1	687	13	AZ733984
25	151.8	12.7	700	11	BE65008
26	148.2	12.4	673	11	BE921319
27	146.4	11.7	640	10	BE289740
28	135.8	11.3	938	10	AL552902
29	133.4	11.1	615	11	BF313287
30	133	11.1	616	11	BE611254
31	132.4	11.1	790	10	AW133824
32	131.8	11.0	680	10	AW077960
33	130.2	10.9	695	11	BE911051
34	129.4	10.8	879	10	AL553876
35	125.2	10.5	467	13	AQ388418
36	119.8	10.0	795	10	AL561703
37	118.8	9.9	467	10	AI218191
38	117	9.8	591	10	AW077347
39	116	9.7	790	11	BE685195
40	115.6	9.7	481	10	AI892564
41	114	9.5	715	10	AI588721
42	113.6	9.5	957	10	AL582021
43	110.8	9.3	532	10	AW182856
44	110	9.2	1068	13	CNS03YB
45	109	9.1	810	10	AI958899

#### ALIGNMENTS

RESULT 1  
BI181178/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI181178 757 bp mRNA EST 10-JUL-2001  
UNL-P-FN-ad-b-04-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone  
UNL-P-FN-ad-b-04-0-UNL 3', mRNA sequence.

BI181178.1 GI:14655587

EST.

Sus scrofa

pig.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

(bases 1 to 757)

Caetano, A.R., Johnson, R.K., and Pomp, D.

Generation and sequence characterization of a normalized cDNA

library from swine ovarian follicles

Unpublished (2001)

Contact: Pomp, D

Department of Animal Science

University of Nebraska, Lincoln

Lincoln, NE 68583-0908, USA

Tel: 402 472 6416

Fax: 402 472 6362

Email: dpomp@unl.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. The following repetitive

elements were found in this cDNA sequence: 495-552,

>GC-richLow-complexity

Seq primer: M13 -29

POLYA-No.

Location/Qualifiers

1..757

/organism="Sus scrofa"

/strain="University of Nebraska, Lincoln Swine Selection

Lines"

```

Ddb      568   CGTCTGGTGTTGAGCCTCTGTGCCATTGCTTTAGAGCGCCACCTTACCATGGC --CGTC 624
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy       421   gggccgcgccccgtctcccagtcggggcgcacgctggcagtggacgccggcctggg 478
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Ddb      625   GTGACCCGCACGCCGCCAGTCGCCTCGCACGCTGCANATGGCGGTGGCGCTGGGG 682
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT    4
BI329992
LOCUS     763 bp mRNA EST 30-JUL-2001
GI        602980556F1 NC_IGAP_L19 Mus musculus cDNA clone IMAGE:5133565 5'
DEFINITION mRNA sequence.
ACCESSION BI329992
VERSION   BI329992
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS  1 (bases 1 to 763)
TITLE    NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL  National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT  Unpublished (1999)
CONTACT  Contact: Robert Strausberg, Ph.D.
```

```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L2AM11327 row: p column: 14
High quality sequence stop: 749.
Location/Qualifiers
1. 763
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5133565"
/clone.lib="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT;
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
108 a 252 C 242 g 161 t
BASE COUNT
ORIGIN

```

Query Match	24.5%	Score 292.8	DB 11	Length 763
Best Local Similarity	81.0%	Pred. No. 9.7e-43		
Matches 426	Conservative 0	Mismatches 92	Indels 8	Gaps 7
Qy 1	atggagtcgggctgctgcggcgccgcgcggtgagcgaggtcatctctgcattacaac	60		
Db 241	ATGAGCCCGGGGTGCTGGCGCGCGCGGTGAGTGAGGTATTGTCTTCACTACAAC	300		
Qy 61	tacacgggcaagctcgcggtgctgcgcgtaccacagcggtgccgcgctgcgcccgaacgc	120		
Db 301	TACACGGGCAAGCTCGGCGGGGCGGCTACCGGCCCGCGGCTGGCGCGACAGCGC	360		
Qy 121	g-tgggtgctcgtggcggtgtgcgcttcctgctgtgctagagaactagcgtgtgttgtgt	179		
Db 361	GCAGATGTCTCTGGGTGTGTGCTTCAATGTGCTGAGAACTGGCTGTGCTCTTGGT	420		
Qy 180	gtcggagccacccgcgcttccacgctcccatgtctctcgtcctcgtggcagcctcaactt	239		
Db 421	GCTGGTCCGCCACCCCTCGCTTCCATGCGGCCCATGTTTCTGCTCTTGGCAGGCTCAC	480		
Qy 240	gtcggaatctgtggcaggcgcgctactacgcgcgcaacatctactctgcgggcgctcac	299		
Db 481	GTGGAGCTTGCTCGCGGGGCGGCTTAGCGCCACCACATCTTACTGTCCGGGCGGCTCAC	540		
Qy 300	gctgaaa-ctgtcccccgcgctctggttcgcacaggga-gggaggcgtcttcg-tggcact	356		
Db 541	GCTGGCGGCTCTFCGCGCTCGCTTGGGTTTTCGGGCTGAGGGGGGGGCTCTTCGATGGCGCT	600		
Qy 357	cactgcgtccgtgctgagacctctctggccatcgctcggagc-gcagcctcaccatggcgc	415		
Db 601	CGCCGCTCGGTGTGAGCCTCTTGGCCATGTGTTTAGAGCAGCCACCTTACCATGGGCC	660		
Qy 416	gcaggggcgccgcgcccgtctccagtgcgggggcgcagctggcgtgagcgagcgcggcct	475		
Db 661	GTCTGTGAGCCGACACCCGGGACAGTGCCTCGCACGCTGGCAATGGGGTGGCGGC--	718		
Qy 476	ggggcgctgctgctcctcgggctcgtccagcgcgtgggctggaa	521		
Db 719	TGGGCGCTCGCTANTGCTGGGCT-GTAACCGCGTGGGTTGGAA	763		

RESULT	5				
BF353652/c					
LOCUS		295 bp	mRNA	EST	22-NOV-2000
DEFINITION					
QV2-HT0698-040700-236-d01 HT0698					Homo sapiens cDNA, mRNA sequence.
ACCESSION					
BF353652					
VERSION					
BF353652.1					GI:11312726

RESULT 6











## High quality sequence stop: 789.

FEATURES  
source

Location/Qualifiers  
1. 793  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4849349"  
/clone\_lib="NIH\_MGC\_106"  
/tissue\_type="natural killer cells, cell line"  
/note="Organ: blood; Vector: pOTB7; Site: 1; XhoI; Site: 2; EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
95 a 284 c 270 g 144 t

BASE COUNT  
ORIGIN

Query Match 18.1%; Score 216.6; DB 11; Length 793;  
Best Local Similarity 59.0%; Pred. No. 3.3e-29;  
Matches 391; Conservative 0; Mismatches 269; Indels 3; Gaps 1;

Qy 13 ctgctgcggcgccgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 72  
Db 69 CAGCTGCG 128

Qy 73 ctccgctgctgcctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 132  
Db 129 CTGCG 188

Qy 133 gcgggtgctgcctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 192  
Db 189 GCCGCCAGCTGCG 248

Qy 193 ccgctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 252  
Db 249 ATGCGGTGCG 308

Qy 253 gcaggcgccgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 312  
Db 309 ACGGCG 368

Qy 313 ccgctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 372  
Db 369 CCG 428

Qy 373 agcctctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 432  
Db 429 AGCTGCTTCTACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 488

Qy 433 ---gtctcagctcggcgccgctaccagccgggtgagcaggttcattcgtcattacacac 489  
Db 489 GGGGCCACCAAGACCG 548

Qy 490 ctctcctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 549  
Db 549 CTGCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608

Qy 550 actgtctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 609  
Db 609 AGCTTCTGCG 668

Qy 610 atcctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 669  
Db 669 GTCTGCG 728

Qy 670 cgg 672  
Db 729 CAG 731

RESULT 12  
AL560216

LOCUS AL560216 LTI\_FLO11\_BCI Homo sapiens cDNA clone CS0DG002YK22 5 prime  
DEFINITION mRNA sequence.  
ACCESSION AL560216 GI:12906465  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 886)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

Location/Qualifiers  
1. 886  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DG002YK22"  
/clone\_lib="LTI\_FLO11\_BCI"  
/sex="male"  
/tissue\_type="B cells from Burkitt lymphoma"  
/lab\_host="DH10B"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"  
BASE COUNT 103 a 311 c 299 g 172 t 1 others  
ORIGIN

Query Match 18.1%; Score 216.6; DB 10; Length 886;  
Best Local Similarity 59.0%; Pred. No. 3.4e-29;  
Matches 391; Conservative 0; Mismatches 269; Indels 3; Gaps 1;

Qy 13 ctgctgcggcgccgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 72  
Db 44 CAGCTGCG 103

Qy 73 ctccgctgctgcctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 132  
Db 104 CTGCG 163

Qy 133 gcgggtgctgcctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 192  
Db 164 GCG 223

Qy 193 ccgctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 252  
Db 224 ATGCGGTGCG 283

Qy 253 gcaggcgccgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 312  
Db 284 ACGGCG 343

Qy 313 ccgctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 372  
Db 344 CCG 403

Qy 373 agcctctgctaccagccgggtgagcaggttcattcgtcattacacacacggcgaag 432

```

Db 404 AGCTGCTCTTCACTGAGGGAGCGCTTTGCCACCATGGTGGGCGCGGTGCCGAGACC 463
QY 433 ----gtctccatcgggggcacgcctgggatggcagcgcgagcctggggcgctgtcgtg 489
Db 464 GGGGCCACCAAGACAGCGCGCTCTACGGCTTCATCGGCTCTGCTGGCTGTGCGCCGCG 523
QY 490 ctctctcgggctccctgcagcgcctggctggaattgcctgggtgcctggagccttgcctc 549
Db 524 CTGCTGGGATGCTGCTTTGCTGGGCTGGAACCTGCTGTGGCCCTTTGACCGCTGCTCC 583
QY 550 actgtcttgcgcctctacgccaaggcctacgtgctgtcttcttgcgtgccttctgttgc 609
Db 584 AGCCTTCTGCCCCCTACTCAAGCGCTACATCCTCTTCTGCTGGTGATCTTCGCCGCGC 643
QY 610 atctcgccctactgtgcaactctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 669
Db 644 GTCTGGCCACCATATGGCGCTCTATGGGCGCATCTTCGCGCTGGTGAGGCCAGCGGG 703
QY 670 cgg 672
Db 704 CAG 706

RESULT 13
AL560608 907 bp mRNA EST 16-FEB-2001
LOCUS AL560608 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL003YC21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL560608
VERSION AL560608.1 GI:12907235
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 907)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL003YC21"
/clone_lib="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6; Site: 1. NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 105 a 327 c 299 g 176 t
ORIGIN

Query Match 18.0%; Score 215.2; DB 10; Length 907;
Best Local Similarity 56.3%; Pred. No. 6.1e-29;
Matches 508; Conservative 0; Mismatches 353; Indels 41; Gaps 4;

QY 13 ctgctcggccggcgccggtagagtgatcgtctcttcattacactacacggcaag 72
Db 29 CAGCTGGCGCGCGGGGACAGCGCGCTCATTTGTTGCTACTACACCACTCGGGCGG 88

```

```

QY 73 ctccgcggtgcgcgtaccagccgggtgcggcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 132
Db 89 CTGGCGGCGCGGGGGGCGC-GGAGGATGGCGGCTGGGGGCGCTTCGGGTG 147
QY 133 ggggtgcgccttccttcgttagaatactagcgcgtgttgggtgcctcggagcgcac 192
Db 148 GCGCCAGCTGCTGGTGTGCTGGAGAACTTGGCTGGTGGCGGCCCATCACCAGCCAC 207
QY 193 cgcgcgttcacgcgtcccatgttccctgcctgggagcgcctcgcgttcgagatctgctg 252
Db 208 ATCGGCTCGCGAGCGCTGGGTCTACATTGCTGGTGGTGAACATCACGCTGAGTGACCTGCTC 267
QY 253 gaaggcgcgcctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 312
Db 268 ACGGGCGCGGCGCTTACCTGGCCCAACGTGCTGTGGGGGGCGCGCACCTTCCTGCTGGCG 327
QY 313 cgcgcgcctcgtgttcgcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 372
Db 328 CCGCGCCAGTGGTCTTAGGGGAGGGCGCTCTTACCGCGCGCTGGCGCGCTTCACCTTC 387
QY 373 agcctcctggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 432
Db 388 AGCCTGCTCTTCACTGACAGGGAGCGCTTTGCCACCATGGTGGCGCGGTGCCGAGAGC 447
QY 433 ---gtctccagtcgggggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 489
Db 448 GGGGCCACCAAGACAGCGCGCGCTTACGGCTCTATCGGCTCTGCTGGCTGCTGGCGCG 507
QY 490 ctctcgggctcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 549
Db 508 CTGCTGGGATGCTGCTTGTGGGCTGGAACTGCTTGTGGCGCTTTGACCGCTTTGACCGCTGCTC 567
QY 550 actgtcttgcgcctctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 609
Db 568 AGCCTTCTGCCCCCTACTTCCAAAGCGCTACATCCTCTTCTGCTGGTGGTATCTTCCCGCGC 627
QY 610 atctcggcgcctctgtgcactctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 669
Db 628 GTCTGGCGCACCATCATGCGGCTCTATGGGCGCATCTTCCCGCGCTGGTGGCGAGCGCGG 687
QY 670 cggcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 729
Db 688 CAGAAGGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
QY 730 cgcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 789
Db 727 -----GCCTGCTGAAGACGCTGCTGATGATCTGCTGGCTTCTTCTGCTGGTGTG 773
QY 790 tggggccccccttctcctcgtctgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 849
Db 774 TGGGCGCCACATCTTGGGCGCTGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 833
QY 850 gtaacctcgcaggcc---gataccttctcggagtgccatgcccactcactcttgaac 906
Db 834 GAGTACCTCGCGGCGCATGGACTGGATCCTGGCCCTGCGCGCTCTCTCAACTCGGGCGGTCAAC 893
QY 907 cc 908
Db 894 CC 895

RESULT 14
BF281109 745 bp mRNA EST 28-NOV-2000
LOCUS BF281109
DEFINITION EST445700 Rat Gene Index, normalized rat, Rattus norvegicus CDNA
ACCESSION Rattus norvegicus CDNA clone RG1AA02, mRNA sequence.
VERSION BF281109
KEYWORDS BF281109.1 GI:11212179
SOURCE EST.
ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2002, 20:58:03 ; Search time 1761.82 Seconds  
(without alignments)  
11208.366 Million cell updates/sec

Title: US-09-842-316-1

Perfect score: 1197  
Sequence: 1 atggagtcgggctgctg.....cagaaccggctgcagactga 1197

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vl.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.om.\*
- 20: em.or.\*
- 21: em.ov.\*
- 22: em.pat.\*
- 23: em.ph.\*
- 24: em.pl.\*
- 25: em.ro.\*
- 26: em.sts.\*
- 27: em.sy.\*
- 28: em.un.\*
- 29: em.vl.\*
- 30: em.htgo\_hum.\*
- 31: em.htgo\_inv.\*
- 32: em.htgo\_rod.\*
- 33: em.htg\_hum.\*
- 34: em.htg\_inv.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1197	100.0	1197	6	AX147832	AX147832 Sequence
2	1197	100.0	100680	9	AC011461	AC011461 Homo sapi
3	1195.4	99.9	1197	9	AF317676	AF317676 Homo sapi
4	1195	99.8	1198	6	AX138796	AX138796 Sequence
5	774	64.7	1203	10	AF233649	AF233649 Rattus no
6	768	64.2	206533	2	AC073749	AC073749 Mus muscu
7	766.4	64.0	2437	10	BC012232	BC012232 Mus muscu
8	764.4	63.9	2171	10	AF115249	AF115249 Rattus no
9	734.8	61.4	1165	10	AF327535	AF327535 Homo sapi
10	642.8	53.7	62147	2	AC026510	AC026510 Homo sapi
11	294.8	24.6	978	10	AY011720	AY011720 Lemur cat
12	289.8	24.2	978	10	AY011704	AY011704 Muscardin
13	277.2	23.2	977	4	AY011726	AY011726 Artibeus
14	275	23.0	978	4	AY011694	AY011694 Sorex ara
15	274.6	22.9	978	4	AY011717	AY011717 Ochotona
16	271.4	22.7	978	10	AY011710	AY011710 Dipodomys
17	271.2	22.7	1149	10	AF108019	AF108019 Mus muscu
18	269.6	22.5	1149	6	AX085541	AX085541 Sequence
19	269.6	22.5	7177	10	W6040811	U40811 Mus musculu
20	268.4	22.4	370	6	AX147764	AX147764 Sequence
21	268.2	22.4	978	10	AY011708	AY011708 Hystric b
22	268	22.4	978	10	AY011714	AY011714 Myocastor
23	265.4	22.2	978	4	AY011693	AY011693 Condylura
24	261.6	21.9	2232	6	RN010303	U10303 Rattus norv
25	261.6	21.9	2232	6	AR027719	AR027719 Sequence
26	261.6	21.9	2232	6	I32245	I32245 Sequence 3
27	261.2	21.8	978	10	AF011706	AF011706 Mus muscu
28	259.2	21.7	1695	5	AF164114	AF164114 Eugu rubr
29	258.6	21.6	978	10	AY011703	AY011703 Castor ca
30	257.4	21.5	975	4	AY011739	AY011739 Leopardus
31	257.4	21.5	975	4	AY011740	AY011740 Panthera
32	257.4	21.5	978	4	AY011737	AY011737 Tapirus i
33	257.4	21.5	978	9	AY011718	AY011718 Cynoceph
34	256.8	21.5	978	4	AY011719	AY011719 Tupia mi
35	255.8	21.4	975	4	AY011738	AY011738 Felis cat
36	255.8	21.4	978	4	AY011736	AY011736 Caracotha
37	255.8	21.4	978	10	AY011715	AY011715 Agouti ta
38	254.8	21.3	978	4	AY011735	AY011735 Equus cab
39	254.2	21.2	978	4	AY011729	AY011729 Megaptera
40	253	21.1	978	10	AY011711	AY011711 Heterocep
41	252	21.1	1149	9	AF233365	AF233365 Homo sapi
42	252	21.1	77678	2	AL161741	AL161741 Homo sapi
43	252	21.1	132438	9	HS0575N6	AL109741 Human DNA
44	251.6	21.0	978	4	AY011742	AY011742 Manis pen
45	251	21.0	978	4	AY011732	AY011732 Tragelaph

ALIGNMENTS

RESULT 1  
AX147832  
LOCUS AX147832 1197 bp DNA  
DEFINITION Sequence 77 from Patent WO0136473.  
ACCESSION AX147832  
VERSION AX147832.1 GI:14346837  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1197)  
Vogeli, G., Wood, L.S., Parodi, L.A., Hiebsch, R.R., Lind, P.,  
Slignton, J., Schellin, K.A., Kaytes, P.S., Bannigan, C.M., Ruff, V.,  
Sejltitz, T. and Huff, R.M.  
Novel g protein-coupled receptors  
Patent: WO 0136473-A 77 25-MAY-2001,  
PHARMACIA & UPJOHN COMPANY (US)

REFERENCE  
AUTHORS

TITLE  
JOURNAL

FEATURES  
source  
1. 1197



















\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1	655: contig of 655 bp in length
666	765: gap of 100 bp
766	1456: contig of 691 bp in length
1457	1556: gap of 100 bp
1557	2243: contig of 687 bp in length
2244	2343: gap of 100 bp
2344	3029: contig of 686 bp in length
3030	3123: gap of 100 bp
3130	3785: contig of 636 bp in length
3786	3885: gap of 100 bp
3886	4562: contig of 677 bp in length
4563	4662: gap of 100 bp
4663	5350: contig of 688 bp in length
5351	5450: gap of 100 bp
5451	6133: contig of 683 bp in length
6134	6233: gap of 100 bp
6234	6914: contig of 681 bp in length
6915	7014: gap of 100 bp
7015	7719: contig of 705 bp in length
7720	7819: gap of 100 bp
7820	8523: contig of 704 bp in length
8524	8623: gap of 100 bp
8624	9322: contig of 699 bp in length
9323	9422: gap of 100 bp
9423	10104: contig of 682 bp in length
10105	10204: gap of 100 bp
10205	10897: contig of 693 bp in length
10898	10997: gap of 100 bp
10998	11698: contig of 701 bp in length
11699	11798: gap of 100 bp
11799	12488: contig of 690 bp in length
12489	12588: gap of 100 bp
12589	13272: contig of 684 bp in length
13273	13722: gap of 100 bp
13732	14050: contig of 678 bp in length
14051	14150: gap of 100 bp
14151	14855: contig of 705 bp in length
14856	14955: gap of 100 bp
14956	15656: contig of 701 bp in length
15657	15756: gap of 100 bp
15757	16458: contig of 702 bp in length
16459	16598: gap of 100 bp
16599	17250: contig of 692 bp in length
17251	17350: gap of 100 bp
17351	18034: contig of 684 bp in length
18035	18134: gap of 100 bp
18135	18814: contig of 680 bp in length
18815	18914: gap of 100 bp
18915	19604: contig of 690 bp in length
19605	19704: gap of 100 bp
19705	20365: contig of 661 bp in length
20366	20465: gap of 100 bp
20466	21154: contig of 689 bp in length
21155	21254: gap of 100 bp
21255	21935: contig of 681 bp in length
21936	22035: gap of 100 bp
22036	22721: contig of 686 bp in length
22722	22821: gap of 100 bp
22822	23505: contig of 684 bp in length
23506	23605: gap of 100 bp
23606	24294: contig of 689 bp in length
24295	24394: gap of 100 bp
24395	25093: contig of 699 bp in length
25094	25193: gap of 100 bp
25194	25899: contig of 706 bp in length



Cancer Institute-FCRDC, P.O. Box B, Bldg. 560, Rm. 11-10,  
Frederick, MD 21702, USA

[illegible]











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2002, 21:03:03 ; Search time 69.25 seconds  
(without alignments)  
3914.717 Million cell updates/sec

Title: US-09-842-316-1

Perfect score: 1197

Sequence: 1 atggagtcgggctgtctgcg.....cagaaccggtgcagactga 1197

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261.6	21.9	2232	1	US-08-196-989B-3
2	261.6	21.9	2232	2	US-08-760-936-3
3	248.4	20.8	1137	3	US-09-082-088-1
4	244	20.4	1637	3	US-08-852-824-3
5	227.4	19.0	1649	2	US-08-845-566-2
6	206.6	17.3	2754	1	US-08-196-989B-1
7	206.6	17.3	2754	2	US-08-760-936-1
8	148.6	12.4	1260	3	US-08-789-982-1
9	137.8	11.5	1761	3	US-08-861-747-1
10	137.8	11.5	1889	3	US-08-861-747-3
11	133.2	11.1	2250	3	US-08-763-938-1
12	98.8	8.3	1875	5	PCT-US93-10618-1
13	90	7.5	2185	2	US-08-467-948A-3
14	90	7.5	2185	3	US-08-467-947A-3
15	67.6	5.6	1417	4	US-09-199-737-3
16	67.2	5.6	1065	4	US-09-325-897-1
17	67	5.6	1356	3	US-08-997-803-13
18	66.4	5.5	1523	3	US-08-997-803-12
19	66.4	5.5	1621	1	US-08-722-001-13
20	66.4	5.5	2140	1	US-08-334-698-1
21	66.4	5.5	2140	1	US-08-228-932-1
22	66.4	5.5	2140	2	US-08-468-939-1
23	66.4	5.5	2140	2	US-08-406-855A-1
24	66.4	5.5	2140	2	US-08-722-190-1
25	66.4	5.5	2140	3	US-08-244-354-1
26	66.4	5.5	2140	3	US-09-206-899-1
27	66.4	5.5	2140	5	PCT-US95-04203-1

28	64.8	5.4	1776	1	US-08-722-001-29	Sequence 29, Appl
29	64	5.3	639	3	US-08-997-803-3	Sequence 3, Appl
30	63	5.3	2580	3	US-09-050-863-2	Sequence 2, Appl
31	63	5.3	5452	2	US-09-130-114-1	Sequence 1, Appl
32	63	5.3	9600	4	US-08-910-647-1	Sequence 1, Appl
33	63	5.3	10596	1	US-07-884-811-15	Sequence 15, Appl
34	63	5.3	10596	1	US-07-885-971-15	Sequence 15, Appl
35	63	5.3	10596	1	US-08-087-783A-15	Sequence 15, Appl
36	63	5.3	10596	2	US-08-194-088B-15	Sequence 15, Appl
37	63	5.3	10596	2	US-08-194-087-15	Sequence 15, Appl
38	63	5.3	10596	5	PCT-US93-04648-15	Sequence 15, Appl
39	62.2	5.2	1296	1	US-07-816-283-9	Sequence 9, Appl
40	62.2	5.2	1296	1	US-08-417-103-9	Sequence 9, Appl
41	59.4	5.0	2126	2	US-08-789-354-1	Sequence 1, Appl
42	59.4	5.0	2126	3	US-09-110-937-1	Sequence 1, Appl
43	59.4	5.0	2126	3	US-09-058-725B-1	Sequence 1, Appl
44	59.4	5.0	2126	3	US-09-232-857-1	Sequence 1, Appl
45	58.4	4.9	1394	2	US-08-068-729-3	Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-08-196-989B-3  
; Sequence 3, Application US/08196989B  
; Patent No. 5585476  
; GENERAL INFORMATION:  
; APPLICANT: MacLennan, A. John  
; TITLE OF INVENTION: Molecular Cloning and Expression of  
; TITLE OF INVENTION: G-Protein Coupled Receptors  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/196,989B  
; FILING DATE: 15-FEB-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lloyd, Jeff  
; REGISTRATION NUMBER: 35,589  
; REFERENCE/DOCKET NUMBER: MAC-100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2232 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 269..1420  
US-08-196-989B-3

Query Match 21.9%; Score 261.6; DB 1; Length 2232;  
Best Local Similarity 56.4%; Pred. NO. 1.7e-40;  
Matches 536; Conservative 0; Mismatches 399; Indels 15; Gaps 2;

OY 37 gaggtcatcgcttcgtacatacaccgcaagtcggtgcgtaccagccg 96









[illegible]



[illegible][illegible]

RESULT 9  
US-08-861-747-1  
; Sequence 1, Application US/08861747  
; Patent No. 6020158  
; GENERAL INFORMATION:  
; APPLICANT: MUNROE, Donald G.  
; APPLICANT: VYAS, Tejal B.  
; TITLE OF INVENTION: A HUMAN EDG-6 RECEPTOR HOMOLOG  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Nikaïdo, Marmelstein, Murray & Oram LLP  
; STREET: 655 15th St., NW, Suite 330 - G Street Lobby  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/861,747  
; FILING DATE: 22-MAY-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jahns, Kristina M.  
; REGISTRATION NUMBER: 41,092

```

; ; REFERENCE/DOCKET NUMBER: P8074-70003
; ;
; ; TELECOMMUNICATION INFORMATION:
; ;
; ; TELEPHONE: (202) 638-5000
; ;
; ; TELEFAX: (202) 638-4810
; ;
; ; INFORMATION FOR SEQ ID NO: 1:
; ;
; ; SEQUENCE CHARACTERISTICS:
; ;
; ; LENGTH: 1761 base pairs
; ;
; ; TYPE: nucleic acid
; ;
; ; STRANDEDNESS: double
; ;
; ; TOPOLOGY: linear
; ;
; ; MOLECULE TYPE: cDNA
; ;
; ; US-08-861-747-1

```

[illegible]

```

Qy 894 ctaactctgaaccccatcatctacacgctaccacacccgacactgcgcccacgctctc 953
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 951 CTCACGTGTCAAATGCTGCTGTGTACACTCTTCCCGAGATGCTGAGATGCCGCCACCTCCG 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 954 gcgcctggtctctgcggacgcccac 978
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1011 CCGCCTTCTGCTGCGCGTGCCTC 1035
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-08-861-747-3
; Sequence 3, Application US/08861747
; Patent No. 6020158
; GENERAL INFORMATION:
; APPLICANT: MUNROE, Donald G.
; APPLICANT: VIAS, Tejal B.
; TITLE OF INVENTION: A HUMAN EDG-6 RECEPTOR HOMOLOG
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., NW, Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,747
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jahns, Kristina M.
; REGISTRATION NUMBER: 41,092
; REFERENCE/DOCKET NUMBER: P8074-7003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1889 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-861-747-3

```

Query Match	11.5%;	Score 137.8;	DB 3;	Length 1889;
Best Local Similarity	49.4%;	Pred. No. 1.1e-17;		
Matches 437;	Conservative 0;	Mismatches 147;	Indels 21;	Gaps 3
QY 54	ttacaactacacgcgcgaagctccgcggtgcgcgtaccagccgggtgcggcctgcgcgcg	113		
Db 218	TTTCTTCTATAACACAGTGGCAAGAGCTCAGCTCCCACTGGCGGCCAAGGATGTGGT	277		
QY 114	cgacgcgtggtgtgcctggtggcgtgtgcgccttcctgctctagagatactagccgtgtt	173		
Db 278	CGTGGTGGGCACATGGGGCTGACCGTGCAGCGCTGTGGTGTGCTGACCAATCTGTGTGGTTCAT	337		
QY 174	gttgggtgcgcggacgccaccgcgcgtttccacgcgtcccatgttctctgctctggcagcct	233		
Db 338	AGCAGCCATCGCCTCCAACCGCCGGTTCACAGCCCATCTACTACCTGTGCTGGCAATCT	397		
QY 234	caagtgttcggatctgtctggcaggcgcgcctcctagccgcgaacatcctactgtctcggggcc	293		
Db 398	GGCGCGGGTGACCTCTCTCGGGGGCGTGGCTACTCTTCTCTATGTTCACACTCGTGCC	457		
QY 294	gcacacgctgaaactgtctcccgcgccttggtctgcacggaaggaagcgtctcttcagc	353		

Db	458	CCGCACAGCCGACCTTCTTCACTTGAGGGCTGGTTCTCGCGCAGGGCTTGCTGGACACAAAG	517
Qy	354	actcaactcgctcagtgctgagcctctctggccatctgcgctggagcgagcctcaccatggc	413
Db	518	CCTCAGTCGGTGGTGCCACACTGCTGGCCACTGCGCGTGGAACGGCACCGCAGTGTGAT	577
Qy	414	gcgcagggggccgcgcgcgctccesagtccggggggcgcgcgctggcgatggcagccgcgc	473
Db	578	GGCGGTACAGTTGCACAGCCGCTGCCCGTGGCGGTGGTCATGTGGGGCT	637
Qy	474	ctggggcggttcgctgctcctcctcctcgggtcctgcagcgctgggctggaattgcctgggctcg	533
Db	638	GTGGGTGGCTGCCCTGGGGCTGGGGCTGTGGCTGCCACTCTTGGGCATGCTCTGTGC	697
Qy	534	cttgagcgctgtctccactgtcttcgcgctctacgccaaaggcctacgctgctctctcgct	593
Db	698	CTTGGACCGTGTCTACGGCATGGCACCCCTGCTCAGCCGCTCTATTTTGGCGGTCTGGGC	757
Qy	594	gctcgctctcgttggggcatccttggcgctatctgtgcactctacgcgcgcactactcgca	653
Db	758	TCTGTGAGCGCTGTTGTCTTCCTGCTCATGGTGGCTGTGTACACCCGCAATTTTTTATA	817
Qy	654	ggtaacggccaaacgcgcgcgctccgcgcacgcgccgggagctggggggacacactcgac	713
Db	818	CGTGGCGGGCGAGTGCACGGCATGCGAGGAT-----GTACAGTGGCCACCCCGGTAC	872
Qy	714	ccggggcgctgcgaacgcgcgcgctcgtggccttgcgcgcgcctcagcgtgggtgcct	773
Db	873	CGAG-----AGACCACGCTCAGCTGCTGAAGACTGTTGTTCATCATCCTGGG	919
Qy	774	ggccttttggcgatgttggggccccccttctcgtcgtgttgcctgacgctggcgctgcc	833
Db	920	GGCGTTCTGTGGTCTGTGACACACAGGCCAGGTGTACTGCTCTGTGATGGTTTAGGCTG	979
Qy	834	ggcgcacactgtcctgtactcctcagggcgcgatcccttctcctggagctggccatggccaa	893
Db	980	TGAG---TCTGCAATGCTCTGGCTGTAGAAAGTACTTCTACTGTTGGCCGAGGCCAA	1036
Qy	894	ctcaacttgaaccccatctcacgctcaccaacgcgcgcgcactggcgcacgcgctcct	953
Db	1037	CTCAGTGGTCATGTGCTGTGTACTCTTGGCGAGATGCTGAGATGCGCGGACACTTCCG	1096
Qy	954	gcgcctggctcgtcggaacgcac	978
Db	1097	CGCCCTCTCTGCTGCGGTGCTC	1121

RESULT 11

```

000001 11
000002 US-08763-938-1
000003 ; Sequence 1, Application US/08763938
000004 ; Patent No. 6140060
000005 ; GENERAL INFORMATION:
000006 ; APPLICANT: CHUN, Jerold J.M.
000007 ; APPLICANT: HECHT, Jonathan H.
000008 ; TITLE OF INVENTION: CLONED LYSOPHOSPHATIDIC ACID
000009 ; TITLE OF INVENTION: RECEPTORS
000010 ; NUMBER OF SEQUENCES: 6
000011 ; CORRESPONDENCE ADDRESS:
000012 ; ADDRESSEE: Nikaïdo, Marmelstein, Murray and Oram LLP
000013 ; STREET: 655 15th Street, N.W., Suite 330 - G St. Lobby
000014 ; CITY: Washington
000015 ; STATE: DC
000016 ; COUNTRY: USA
000017 ; Zip: 20005-5701
000018 ; COMPUTER READABLE FORM:
000019 ; MEDIUM TYPE: Floppy disk
000020 ; COMPUTER: IBM PC compatible
000021 ; OPERATING SYSTEM: PC-DOS/MS-DOS
000022 ; SOFTWARE: Patentin Release #1.0, Version #1.30
000023 ; CURRENT APPLICATION DATA:
000024 ; FILING DATE: 12-DEC-1996
000025 ; APPLICATION NUMBER: US/08/763,938

```











GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 23:38:08 ; Search time 112.43 Seconds  
 (without alignments)  
 517.801 Million cell updates/sec

Title: US-09-842-316-2  
 Perfect score: 2019  
 Sequence: 1 MESGLLRPAPEVSEVILVHYN.....TGSPGAPTARTLVSEPAAD 398

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
 Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SPTREMBL\_17.\*  
 1: sp-archaea.\*  
 2: sp-bacteria.\*  
 3: sp-fungi.\*  
 4: sp-human.\*  
 5: sp-invertebrate.\*  
 6: sp-mammal.\*  
 7: sp-mhc.\*  
 8: sp-organella.\*  
 9: sp-phage.\*  
 10: sp-plant.\*  
 11: sp-rodent.\*  
 12: sp-virus.\*  
 13: sp-vertebrate.\*  
 14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2019	100.0	398	4	Q9H228
2	1714.5	84.9	400	11	Q9JKM5
3	1698.5	84.1	400	11	Q9QY79
4	1625.5	80.5	387	11	Q9NMN8
5	865.5	42.9	382	11	Q9R235
6	863.5	42.8	382	11	Q9DC35
7	838.5	41.5	382	4	Q9NYN8
8	814.5	40.3	362	13	Q9DDK4
9	814	40.3	326	11	Q9NR55
10	812	40.2	326	11	Q9NR88
11	805	39.9	326	6	Q9BF53
12	804	39.8	326	6	Q9BF72
13	804	39.8	326	11	Q9NR44
14	803	39.8	326	6	Q9BF60
15	803	39.8	326	11	Q9NQ77
16	802	39.7	326	6	Q9BF73
17	802	39.7	326	6	Q9BF46
18	802	39.7	326	11	Q9NQ99
19	801	39.7	326	11	Q9NR11

Q9bf75 myrmecophag  
 Q99nr0 cavia tschu  
 Q9bf79 choloepus h  
 Q9bf78 euphractus  
 Q9bf45 ceratotheri  
 Q9bf44 tapirus ind  
 Q99nr7 muscardinus  
 Q9bf42 leopardus p  
 Q9bf41 panthera on  
 Q9bf76 tamandua te  
 Q9bf74 erinaceus c  
 Q9bf62 cynocephalu  
 Q99nr2 dipodomys h  
 Q9bf67 macroselid  
 Q99nr3 erethizon d  
 Q9bf69 procavia ca  
 Q9bf65 orycteropus  
 Q99nr9 tamias stri  
 Q9bf70 trichechus  
 Q9bf68 loxodonta a  
 Q9byy4 homo sapien  
 Q9bf57 hylobates c  
 Q9bf77 chaetophrac  
 Q9bf66 elephantulu  
 Q9bf58 macaca mula  
 Q9puq8 fugu rubrip

## ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	398 AA.
Q9H228	Q9H228			
AC	Q9H228			
DT	01-MAR-2001 (TRENBLrel. 16, Created)			
DT	01-MAR-2001 (TRENBLrel. 16, Last sequence update)			
DE	SPHINGOSINE 1-PHOSPHATE RECEPTOR EDG-8.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Im D., Lynch K.R.;			
RT	"Characterization of a human sphingosine 1-phosphate receptor, Edg-8."			
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF317676; AAG38113.1;			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCRHHODPSN.			
DR	PROSITE; PS00237; G-PROTEIN_RECEP_FL1; UNKNOWN_1.			
DR	PROSITE; PS00262; G-PROTEIN_RECEP_FL2; 1.			
KW	Receptor.			
SQ	SEQUENCE 398 AA; 41775 MW; 34ADC6C40D8250BF CRC64;			

Query Match 100.0%; Score 2019; DB 4; Length 398;  
 Best Local Similarity 100.0%; Pred. No. 5,1e-125;  
 Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MESGLLRPAPEVSEVILVHYN	TGKRGARYOPAGLRADAVVCLAVCAFTVLENVAVLV 60
DB	1	MESGLLRPAPEVSEVILVHYN <td>TGKRGARYOPAGLRADAVVCLAVCAFTVLENVAVLV 60</td>	TGKRGARYOPAGLRADAVVCLAVCAFTVLENVAVLV 60
QY	61	LCRHPRFAPFLLIGSLTSLDLAAYAAANTLLSGPLTKLSPALWFAREGGVVAVLT 120	
DB	61	LCRHPRFAPFLLIGSLTSLDLAAYAAANTLLSGPLTKLSPALWFAREGGVVAVLT 120	
QY	121	ASVLSLAIALERSLTWARRGPAPVSRGRTLMAAAWGVSLLLGLPALGVNCLGRD 180	

```
Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLD 180
Qy 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGTACTTSTRA 240
Db 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGTACTTSTRA 240
Qy 241 RRRPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 300
Db 241 RRRPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 300
Qy 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPGLDGSF 360
Db 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPGLDGSF 360
Qy 361 SGERSSPQRDGLDTSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLDGSF 398
Db 361 SGERSSPQRDGLDTSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLDGSF 398

RESULT 2
Q9JKM5 PRELIMINARY; PRT; 400 AA.
ID Q9JKM5 AC Q9JKM5
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SPHINGOSINE 1-PHOSPHATE RECEPTOR EDG-8.
EN EDG-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20261551; PubMed=10799507;
RA Im D.-S., Heise C.E., Ancellin N., O'Dowd B.F., Shen G.-J.,
RA Heavens R.P., Rigby M.R., Hla T., Mandala S., McAllister G.,
RA George S.R., Lynch K.R.;
RT "Characterization of a novel sphingosine 1-phosphate receptor, Edg-
RT 8."
RL J. Biol. Chem. 275:14281-14286(2000).
DR EMBL; AF233649; AAF35912.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 400 AA; 42368 MW; 71B398DC4DE0FEB CRC64;

Query Match 84.9%; Score 1714.5; DB 11; Length 400;
Best Local Similarity 87.2%; Pred. No. 4.6e-105;
Matches 346; Conservative 14; Mismatches 34; Indels 3; Gaps 3;

Qy 1 MESGLLRPAVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
Db 1 MESGLLRPAVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
Qy 61 LGRHPRFHAPMFLGLSLTSLDLAAGAAAYANILLSGPLTKLSPALWFAREGGVFALT 120
Db 61 LGRHPRFHAPMFLGLSLTSLDLAAGAAAYATNILLSGPLTKLSPALWFAREGGVFALA 120
Qy 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLD 180
Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLE 180
Qy 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGT-AGTTSTR 239
Db 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGTSSSR 240
Qy 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 299
Db 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 299
Qy 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLD 180
Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLE 180
Qy 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGT-AGTTSTR 239
Db 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGTSSSR 240
Qy 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 299
Db 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 299
```

```
Db 241 SRHTPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 300
Qy 300 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPGLD 357
Db 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPGLD 360
Qy 358 GSFSSERSPPQRDGLDTSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLD 394
Db 361 RSSPSERHSCQRDGMDSCTSGSPGAATANTLVPD 397

RESULT 3
Q9QY79 PRELIMINARY; PRT; 400 AA.
ID Q9QY79 AC Q9QY79
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GROWTH FACTOR-REGULATED G PROTEIN-COUPLED RECEPTOR NRG-1.
EN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99453967; PubMed=10532805;
RA Gluckman M., Malek R.L., Wittek-Black A.E., Jacob H.J., Lee N.H.;
RT "Molecular cloning, tissue-specific expression, and chromosomal
RT localization of a novel nerve growth factor-regulated G-protein-
RT coupled receptor, nrg-1."
RL Mol. Cell. Neurosci. 14:141-152(1999).
DR EMBL; AF115249; AAF15395.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 400 AA; 42368 MW; 72B46CCE3BD24D4B CRC64;

Query Match 84.1%; Score 1698.5; DB 11; Length 400;
Best Local Similarity 86.1%; Pred. No. 5.1e-104;
Matches 342; Conservative 16; Mismatches 36; Indels 3; Gaps 3;

Qy 1 MESGLLRPAVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
Db 1 MESGLLRPAVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
Qy 61 LGRHPRFHAPMFLGLSLTSLDLAAGAAAYANILLSGPLTKLSPALWFAREGGVFALT 120
Db 61 LGRHPRFHAPMFLGLSLTSLDLAAGAAAYATNILLSGPLTKLSPALWFAREGGVFALA 120
Qy 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLD 180
Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLGLLGLPALGWNCLGRLE 180
Qy 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGT-AGTTSTR 239
Db 181 ACSTVPLXAKAYVFCVLAFLAFVILAAICALYARIYCOVRANARLPARPGTSSSR 240
Qy 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 299
Db 241 SRHTPRSLALLRTLSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLAMANS 300
Qy 300 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPGLD 357
Db 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPGLD 360
Qy 358 GSFSSERSPPQRDGLDTSVLLAFVACWGPFLFLLLLDLDVACPARTCPVLLQADPFLGLD 394
Db 361 RSSPSERHSCQRDGMDSCTSGSPGAATANTLVPD 397
```

```

RESULT 4
Q99MN8 PRELIMINARY; PRT; 387 AA.
ID Q99MN8
AC Q99MN8
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE SPHINGOSINE-1-PHOSPHATE RECEPTOR LPB4 (FRAGMENT).
GN LPB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER/NIH;
RA Yang A.H., Zhang G., Chun J.J.M.;
RT "Molecular cloning of the mouse sphingosine-1-phosphate receptor gene,
RT Lpb4."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327535; AAK15485.1;
KW NON_TER
FT
SQ SEQUENCE 387 AA; 40953 MW; B56B47672BF5E977 CRC64;

Query Match 80.5%; Score 1625.5; DB 11; Length 387;
Best Local Similarity 86.1%; Pred. No. 3e-99;
Matches 329; Conservative 14; Mismatches 36; Indels 3; Gaps 3;

QY 14 VIVLHYNTGKLGARYQPGAGLRADAVVCLVCAFIYVLENLAVLVLRHPRFHAPMFL 73
Db 1 VIVLHYNTGKLGARYQPGAGLRADAVVCLVCAFIYVLENLAVLVLRHPRFHAPMFL 60
QY 74 LLGSLTSLDLAAGAAATNLLSGPLTKLSPALWFAREGGVFVALTASVLSLAIALER 133
Db 61 LLGSLTSLDLAAGAAATNLLSGPLTKLSPALWFAREGGVFVALTASVLSLAIALER 120
QY 134 SLTMARGPAPVSSRRTLMAAAAGVSLGLLPGALWNCGLRDLACSTVLPYAKAY 193
Db 121 HLTMARGPAPASRARTLMAAAAGVSLGLLPGALWNCGLRDLACSTVLPYAKAY 180
QY 194 VLFVCLAFVILAAICALLYARIYCOVRANARLPPGCT-AGTTTRARRRPSRLALLRT 252
Db 181 VLFVCLAFVILAAICALLYARIYCOVRANARLPPGCT-AGTTTRARRRPSRLALLRT 240
QY 253 LSVVLLAFVACMGPLELLLLDLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNDR 312
Db 241 LSVVLLAFVACMGPLELLLLDLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNDR 300
QY 313 LRHALLRLVCCGRHSCGRDPGSG-QQSASAAEAS-GGLRRLCLPPGLDGSFSGSERSSPQR 370
Db 301 LRHALLRLVCCGRHSCGRDPGSG-QQSASAAEAS-GGLRRLCLPPGLDGSFSGSERSSPQR 360
QY 371 DGLDTSGTSGCAPTAARTLV 392
Db 361 DGLDTSGTSGCAPTAARTLV 382

RESULT 5
Q9R235 PRELIMINARY; PRT; 382 AA.
ID Q9R235
AC Q9R235
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE LYOPHOSPHOLIPID RECEPTOR BL.
GN EDGI OR LPB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN
RP SEQUENCE FROM N.A.
RC STRAIN=125SVJ;
RX MEDLINE=9513220; PubMed=9931453;
RA Zhang G., Contos J.J.A., Weiner J.A., Fukushima N., Chun J.;
RT "Comparative analysis of three murine G-protein coupled receptors
RT activated by sphingosine-1-phosphate.";
RL Gene 227:89-99(1999)
DR EMBL; AF108019; AAD16975.1;
DR MGI; MGI:1096355; Edg1.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1.1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1.1; UNKNOWN_1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 382 AA; 42613 MW; 2EB4B974E9FBF39C CRC64;

Query Match 42.9%; Score 865.5; DB 11; Length 382;
Best Local Similarity 48.5%; Pred. No. 1.8e-49;
Matches 101; Conservative 66; Mismatches 105; Indels 21; Gaps 6;

QY 13 EVIVLHYNTGKLR-GARYQPGAGLRADAVVCLVCAFIYVLENLAVLVLRHPRFHAPM 71
Db 23 DIIVRHVNTGKLNIGA--EKDHGKLTSTVVFILCCFIILENIFVLLTIWTKKFRHPM 80
QY 72 ELLGSLTSLDLAAGAAATNLLSGPLTKLSPALWFAREGGVFVALTASVLSLAIAL 131
Db 81 YFIGNLALSDLLAGVAYTANLLSGATTYKLTPAQWFLREGSMFVALSASFVLSLAI 140
QY 132 ERLTMARGPAPVSSRRTLMAAAAGVSLGLLPGALWNCGLRDLACSTVLPYAK 191
Db 141 ERYTLMKMLHNGSNSSRSFLLISACVWISLGLPSMGWNCISSLSGCTVLPYHK 200
QY 192 AYVLCVLAFAVILAAICALLYARIYCOVRANARLPPGCTAGTTTRARRRPSRLALLR 251
Db 201 HYLFCTVFTVLLLSIVILYCRYISLVTRRSRLTFRK----NISKASRSSEKSLALK 256
QY 252 TLSVLLAFVACMGPLELLLLDLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLN 311
Db 257 TVIVLSVLAFAVILAAICALLYARIYCOVRANARLPPGCTAGTTTRARRRPSRLALLR 316
QY 312 DURHALLRLVCCGRHSCGRDPGSG-QQSASAAEAS-GGLRRLCLPPGLDGSFSGSERS-PQR 370
Db 317 EMRRAPFIRIV-----SCCKCPNG-----DSAGKFKRPIIPGNEFSRSKSDNSHPQK 363
QY 371 DGLDTSGTSGSPG 383
Db 364 DGDNDPNTINSSG 376

RESULT 6
Q9DC35 PRELIMINARY; PRT; 382 AA.
ID Q9DC35
AC Q9DC35
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE ENDOTHELIAL DIFFERENTIATION SPHINGOLIPID G-PROTEIN-COUPLED RECEPTOR 1.
GN EDGI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

```

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauble R., Suzuki K., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
DR EMBL: AK004591; BAB23393.1; -;  
DR MGD: MGI:1096355; Edg1.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsin.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
SQ SEQUENCE 382 AA; 42639 MW; 5FE4C9A2BD65CB2A CRC64;

Query Match 42.8%; Score 863.5; DB 11; Length 382;  
Best Local Similarity 48.5%; Pred. No. 2.5e-49;  
Matches 181; Conservative 65; Mismatches 106; Indels 21; Gaps 6;

QY 13 EVIVLHNYTGKLR-GARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHAPMF 71  
Db 23 DIIVRHNYTGKLNIG--EKDHGKLTSVFILICCFIENLAVLLVLRHPRHAPMF 80  
QY 72 FLIGSLTSLDLGAGAAAYANILLGSLTLPALWFAREGGVFVALTSVLSLAIA 131  
Db 81 YFTGNLALSDLAGVAYTANLLSGATYKLTPAQWFLREGSMFVALSVLSLAIAI 140  
QY 132 ERLTHARGPAPVSSRGRTLMAAANGVLLGLLPALGWNCLGRDACSTVLPYAK 191  
Db 141 ERYITLTKMKLHNGSSRSFLLISACWVISLILGLPIMGNCSLSLSSCSTVLPYKH 200  
QY 192 AVVLCVLAFLVGLTAAICALLYARIYCOVRANARLRPARPGTAGTTSRARRKPSLALLR 251  
Db 201 HYILFCTVFTLLLSIVLYCRIYSLVTRSRRTFRK----NISKASRSSEKSLALLK 256  
QY 252 TLSVLLAFVACGPFLLLLLDVACPARTCPVLQADPFLGLAMANSLLNPIIYTLTNR 311  
Db 257 TVIIVLSVFIAWAPLFIILLDVGCKARTCDILYKAEYFLVAVLNSGTNPPIIYTLTK 316  
QY 312 DLRHALLRLVCCGRHSCGRDPSGQSASAAEASGLRCLPPLGDSFGSFSERS--PQR 370  
Db 317 EMRAFIRIV-----SCCKCPNG-----DSAGKFRPIIPGMEFSSRSKSDNSHPQK 363  
QY 371 DGLDTSGETSGSPG 383  
Db 364 DGDNDPNTIMSSG 376

RESULT 7  
Q9NVN8 ID Q9NVN8 PRELIMINARY; PRT; 382 AA.  
AC Q9NVN8  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE G PROTEIN-COUPLED SPHINGOLIPID RECEPTOR.  
GN CHEGDI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Tigyil G.J. Sr., Wang D. Sr.;  
RT "Correction to Human EDG1 sequence";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF233365; AAF43420.1; -;  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsin.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 382 AA; 42810 MW; OCCE8685A5E1BAD2 CRC64;

Query Match 41.5%; Score 838.5; DB 4; Length 382;  
Best Local Similarity 47.3%; Pred. No. 1.1e-47;  
Matches 176; Conservative 66; Mismatches 111; Indels 19; Gaps 5;

QY 13 EVIVLHNYTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHAPMF 72  
Db 23 DIIVRHNYTGKLNISADKENS-IKLTSVFILICCFIENLAVLLVLRHPRHAPMF 81  
QY 73 LILGSLTSLDLGAGAAAYANILLGSLTLPALWFAREGGVFVALTSVLSLAIALE 132  
Db 82 YFTGNLALSDLAGVAYTANLLSGATYKLTPAQWFLREGSMFVALSVLSLAIAIE 141  
QY 133 RLTHARGPAPVSSRGRTLMAAANGVLLGLLPALGWNCLGRDACSTVLPYAKA 192  
Db 142 RYITLTKMKLHNGSNFLLISACWVISLILGLPIMGNCSLSLSSCSTVLPYKH 201  
QY 193 YVLCVLAFLVGLTAAICALLYARIYCOVRANARLRPARPGTAGTTSRARRKPSLALLRT 252  
Db 202 YILFCTVFTLLLSIVLYCRIYSLVTRSRRTFRK----NISKASRSSEKSLALLK 257  
QY 253 LSVLLAFVACGPFLLLLLDVACPARTCPVLQADPFLGLAMANSLLNPIIYTLTNR 312  
Db 258 TVIIVLSVFIAWAPLFIILLDVGCKVKTCDILYKAEYFLVAVLNSGTNPPIIYTLNKE 317  
QY 313 LHAHALLRLVCCGRHSCGRDPSGQSASAAEASGLRCLPPLGDSFGSFSERS--PQR 371  
Db 318 MRRAFIRIV-----SCCKCPNG-----DSAGKFRPIIPGMEFSSRSKSDNSHPQK 364  
QY 372 GLDTSGETSGSPG 383  
Db 365 EGDNDPNTIMSSG 376

RESULT 8  
Q9DDK4 ID Q9DDK4 PRELIMINARY; PRT; 362 AA.  
AC Q9DDK4  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE SPHINGOSINE 1-PHOSPHATE RECEPTOR.  
GN EDG1.  
OS Brachydanio rerio (zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Rasbora; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]

SEQUENCE FROM N.A.  
RP MEDLINE-20563813; PubMed-11112429;  
RX Im D.S., Ungar A.R., Lynch K.R.;  
RT "Characterization of a zebrafish (Danio rerio) sphingosine 1-phosphate  
RT receptor expressed in the embryonic brain";  
RL Biochem. Biophys. Res. Commun. 279:139-143(2000).  
DR EMBL: AF321294; AAC45430.1; -;  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsin.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.

```
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor. 362 AA; 40583 MW; 43A398871DA2D9D4 CRC64;
SQ SEQUENCE 362 AA; 40583 MW; 43A398871DA2D9D4 CRC64;

Query Match 40.3%; Score 814.5; DB 13; Length 362;
Best Local Similarity 47.1%; Pred. No. 3.8e-46;
Matches 172; Conservative 59; Mismatches 107; Indels 27; Gaps 7;

QY 11 VSEVIVHNYTGKLGARYQDAGLRADAVVCLAVCAFIIVLENLAVLLVGRPRHAP 70
DB 1 MDLIARHNYFTGKRVKHKD--GLKADSVYFIIVCCFIILENVLLTWTTRTKKFKP 58
QY 71 MFLLGSLTSLDLAGAAYANILLSPGLTKLSPALWFAREGGVFVLTASVLSLAI 130
DB 59 MYFIGNLALSLLAGVYVYANILLSGANTYKLTPTQWFFREGSMFVALAASVLSLAI 118
QY 131 LERSLTMRARGPAPVSSRGRT---LMAAAAGVSVLLGLLPALGWNCGLRLDACSTVLP 187
DB 119 IEHRLTMLK---MKLHNGKTCRVFMLISTVNFIAAILGGLPVMGNCIDSIINNCSTVLP 175
QY 188 LYAKAVVLCVLAFAVGIILAAICARYICOVANARLRLPARPGTAGTSTRARKPRSL 247
DB 176 LYHKAVILECTVFSVILMAIVILYRIYALVTRSRKLVKRVKANGSKSEK--SM 233
QY 248 ALLRTSVVLLAFVACWGPLFLLLLLVACPARTCPVLLQADPFLGLAMANSLLNPIYT 307
DB 234 ALLKTVIIVLSFCIACWAPFLFLLLLVACOTLCSILYKAEMFLALAVLANSANPLIYT 293
QY 308 LTRNDRHALLRVLCCGRHSCGRDPSGSOASAAEGGLRRLCPGLDGSFSGSERS 367
DB 294 LRSNEMRRAFIAKMLNCG-----VCQPSGKFSRPI-MGAEFSSKSDNS 337
QY 368 -PORD 371
DB 338 HPNKD 342

RESULT 9
Q99NR5 PRELIMINARY; PRT; 326 AA.
AC Q99NR5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011706; AAK01975.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36574 MW; 7B6DF34BD398B1F7 CRC64;

Query Match 40.28; Score 812; DB 11; Length 326;
Best Local Similarity 49.1%; Pred. No. 5.1e-46;
Matches 166; Conservative 61; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLLVGRPRHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GIKLTSVVFILICCFIILENFVLLTWTKTKRPMYFIGNLSDLLAGVAYTANLL 63
QY 95 LSGPLTKLSPALWFAREGGVFVLTASVLSLAIERSLTMRARGPAPVSSRGRTLAM 154
DB 64 LSGATTYKLTQAQFREGSMFVALSASFSLAIATERYITLTKMLHNSNSRSFLL 123
QY 155 AAAAGVSVLLGLLPALGWNCGLRLDACSTVLPYAKAVVLCVLAFAVGIILAAIC 214
DB 124 ISACWVSVLLGLLPALGWNCIGTLSSCSIVPLYKHNYLFTCTVFTLLLSIVLYCR 183
QY 215 IYQVVRANARLRLPARPGTAGTSTRARKPRSLALRTLSVLLAFVACWGPLFLLLL 274
DB 184 IYSLVTRSRRLTRFK---NISKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLL 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIYITNDRHALLRLVCCGRHSCGRDPSG 334
DB 240 VGCKAKTCDILYKAEYFLVAVLANSNGNPIYITLTKEMRRAFIRIV-----SCKCPNG 294
QY 335 SQQSASAAEGGLRRLCPGLDGSFSGSERS--PORD 371
DB 295 -----DSAGKFKRPIPGMEFSSKSDNSHPQKD 324

RESULT 10
Q99NR8 PRELIMINARY; PRT; 326 AA.
AC Q99NR8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Castor canadensis (American beaver).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Castoridae; Castor.
OX NCBI_TaxID=51338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011703; AAK01972.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36552 MW; 4BA3EEC6266D1AFA CRC64;

Query Match 40.28; Score 812; DB 11; Length 326;
Best Local Similarity 49.1%; Pred. No. 5.1e-46;
Matches 166; Conservative 61; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLLVGRPRHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GIKLTSVVFILICCFIILENFVLLTWTKTKRPMYFIGNLSDLLAGVAYTANLL 63
QY 95 LSGPLTKLSPALWFAREGGVFVLTASVLSLAIERSLTMRARGPAPVSSRGRTLAM 154
DB 64 LSGATTYKLTQAQFREGSMFVALSASFSLAIATERYITLTKMLHNSNSRSFLL 123
QY 155 AAAAGVSVLLGLLPALGWNCGLRLDACSTVLPYAKAVVLCVLAFAVGIILAAIC 214
DB 124 ISACWVSVLLGLLPALGWNCIGTLSSCSIVPLYKHNYLFTCTVFTLLLSIVLYCR 183
QY 215 IYQVVRANARLRLPARPGTAGTSTRARKPRSLALRTLSVLLAFVACWGPLFLLLL 274
DB 184 IYSLVTRSRRLTRFK---NISKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLL 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIYITNDRHALLRLVCCGRHSCGRDPSG 334
DB 240 VGCKAKTCDILYKAEYFLVAVLANSNGNPIYITLTKEMRRAFIRIV-----SCKCPNG 294
QY 335 SQQSASAAEGGLRRLCPGLDGSFSGSERS--PORD 371
DB 295 -----DSAGKFKRPIPGMEFSSKSDNSHPQKD 324

Query Match 40.3%; Score 814; DB 11; Length 326;
Best Local Similarity 49.1%; Pred. No. 3.8e-46;
Matches 166; Conservative 61; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLLVGRPRHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GIKLTSVVFILICCFIILENFVLLTWTKTKRPMYFIGNLSDLLAGVAYTANLL 63
QY 95 LSGPLTKLSPALWFAREGGVFVLTASVLSLAIERSLTMRARGPAPVSSRGRTLAM 154
```



```

RESULT 11
Q9BF63 PRELIMINARY; PRT; 326 AA.
AC Q9BF63;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN
OS Ochozona hyperborea (northern pika).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Ochotonidae; Ochotona.
OX NCBI_TaxID=130834;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011717; AAK01985.1; -
FT NON_TER 1 326
FT NON_TER 326
SQ SEQUENCE 326 AA; 36468 MW; 2DEC66EF9439797 CRC64;

Query Match 39.9%; Score 805; DB 6; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.5e-45;
Matches 165; Conservative 62; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVFVILICCFIILENIFVLLTWKTKFHRPMYFIGNLSDLLAGVAYIANLL 63

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSLTARRGPAPVSSRGRTLAM 154
DB 64 LSGATYKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNFSRFL 123

QY 155 AAAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAIYAR 214
DB 124 ISACKVTKTCAILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSLTARRGPAPVSSRGRTLAM 154
DB 64 LSGATYKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNFSRFL 123

QY 155 AAAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAIYAR 214
DB 124 ISACKVTKTCAILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 215 IYQVNRANRLPARPGTAGTSTARRKPRSLALLRSLVLLAFVACWGPLFLLLLD 274
DB 184 IYSLVTRSRRLTFRK---NASKASRSSEKSLALLTVIIVLSVFIACWAPLFI 239

QY 275 VACPARTCVLLQADPFGLAMANSLLNPITYTLNRDLRALRLVCCGRHSCGRDPSG 334
DB 240 VGCKVTKTCDILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 335 SQSASAAASGGLRCLPGLDGSFSGSERSS-PQRD 371
DB 295 -----DSAGKVRPIIAGVEFSRSKSDNSHPQRD 324

Query Match 39.9%; Score 805; DB 6; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.5e-45;
Matches 165; Conservative 62; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVFVILICCFIILENIFVLLTWKTKFHRPMYFIGNLSDLLAGVAYIANLL 63

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSLTARRGPAPVSSRGRTLAM 154
DB 64 LSGATYKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNFSRFL 123

QY 155 AAAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAIYAR 214
DB 124 ISACKVTKTCDILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 215 IYQVNRANRLPARPGTAGTSTARRKPRSLALLRSLVLLAFVACWGPLFLLLLD 274
DB 184 IYSLVTRSRRLTFRK---NASKASRSSEKSLALLTVIIVLSVFIACWAPLFI 239

QY 275 VACPARTCVLLQADPFGLAMANSLLNPITYTLNRDLRALRLVCCGRHSCGRDPSG 334
DB 240 VGCKVTKTCDILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 335 SQSASAAASGGLRCLPGLDGSFSGSERSS-PQRD 371
DB 295 -----DSAGKVRPIIAGVEFSRSKSDNSHPQRD 324

RESULT 12
Q9BF72 PRELIMINARY; PRT; 326 AA.
AC Q9BF72;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN
OS Sorex araneus (Eurasian common shrew) (European shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Sorex.
OX NCBI_TaxID=42254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;

```

```

RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011694; AAK01963.1; -
FT NON_TER 1 326
FT NON_TER 326
SQ SEQUENCE 326 AA; 36577 MW; 5BBB6812E3F635B7 CRC64;

Query Match 39.8%; Score 804; DB 6; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.7e-45;
Matches 165; Conservative 60; Mismatches 95; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVFVILICCFIILENIFVLLTWKTKFHRPMYFIGNLSDLLAGVAYIANLL 63

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSLTARRGPAPVSSRGRTLAM 154
DB 64 LSGATYKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNFSRFL 123

QY 155 AAAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAIYAR 214
DB 124 ISACKVTKTCAILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 215 IYQVNRANRLPARPGTAGTSTARRKPRSLALLRSLVLLAFVACWGPLFLLLLD 274
DB 184 IYSLVTRSRRLTFRK---NASKASRSSEKSLALLTVIIVLSVFIACWAPLFI 239

QY 275 VACPARTCVLLQADPFGLAMANSLLNPITYTLNRDLRALRLVCCGRHSCGRDPSG 334
DB 240 VGCKVTKTCAILFKAIEYFLVAVLNSATNPITYTLTKEMRRAFIRIN-----SCCKCPSG 294

QY 335 SQSASAAASGGLRCLPGLDGSFSGSERSS-PQRD 371
DB 295 -----DSAGKVRPIIAGVEFSRSKSDNSHPQRD 324

RESULT 13
Q99NR4 PRELIMINARY; PRT; 326 AA.
AC Q99NR4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Hystricidae; Hystrix.
OX NCBI_TaxID=143286;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011708; AAK01976.1; -
FT NON_TER 1 326
FT NON_TER 326
SQ SEQUENCE 326 AA; 36546 MW; 3E66CA765F725BA9 CRC64;

Query Match 39.8%; Score 804; DB 11; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.7e-45;
Matches 165; Conservative 60; Mismatches 95; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVFVILICCFIILENIFVLLTWKTKFHRPMYFIGNLSDLLAGVAYIANLL 63

```

```

QY 95 LSGPLTLKSPALMFAREGGVVALTSVLSLALALERSLTMARRGPAPVSSRGRTLAM 154
Db 64 LSGATTYKLTPAQWFLREGSMFALSASFSLAIAIERVITMLKMKLHNGSNSSRSFLL 123
QY 155 AAAAGVSLLLGLLPALGNCNCLGRDCACTVPLPYAKAYVLCVLAFCVLAIAICALYAR 214
Db 124 ISACWVISLILGGLPIMGWNCISLSSCSTVPLPYHKHYILFCTVFTVTLALLAIVLYCR 183
QY 215 IYCOVRANARLPARPAGTAGTSTRARRKPSRLALLRTLSTVLLAFVACWGPLFLLILLD 274
Db 184 IYSLVTRSRRLTFRK ----NISKASRSSEKSLALLKTVIIVLSVFIACWAPLFILLLD 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRDLRHALLRVCCGRHSCGRDPSG 334
Db 240 VGCKVKTCDILFYKAEYFLVAVLNSGTNPITYTLTNKEMRRAFIRIM-----SCKKCPGG 294
QY 335 SQOSASAAEASGGLRCLPPGLDGSFSGSERS-PORD 371
Db 295 -----DSTGKIKRPIIAGVEFSRSKSDNSHPKOD 324

RESULT 14
ID Q9BF60 PRELIMINARY; PRT; 326 AA.
AC Q9BF60;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE EDG1 (FRAGMENT)
GN EDG1.
OS Lemur catta (Ring-tailed lemur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Lemur.
OX NCBI_TaxID=9447;
RN [1]
RP SEQUENCE FROM N.A.
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011720; AAK01988.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36289 MW; B4D31565F3D5A6AD CRC64;

Query Match 39.8%; Score 803; DB 6; Length 326;
Best Local Similarity 49.4%; Pred. No. 2e-45;
Matches 167; Conservative 57; Mismatches 96; Indels 18; Gaps 4;

QY 35 GLRADAVCLAVCAFTVLENLAVLLVGRPHRPHAPMFLLSGLSLDLAGAAYANILL 94
Db 4 GLAPTSPVILICCFIILENIFVLLTWKTKRPHRPMYFIGNLALSOLLAGVAYIANLL 63
QY 95 LSGPLTLKSPALMFAREGGVVALTSVLSLALALERSLTMARRGPAPVSSRGRTLAM 154
Db 64 LSGATTYKLTPAQWFLREGSMFALSASFSLAIAIERVITMLKMKLHNGSNSSRSFLL 123
QY 155 AAAAGVSLLLGLLPALGNCNCLGRDCACTVPLPYAKAYVLCVLAFCVLAIAICALYAR 214
Db 124 ISACWVISLILGGLPIMGWNCISLSSCSTVPLPYHKHYILFCTVFTVTLALLAIVLYCR 183
QY 215 IYCOVRANARLPARPAGTAGTSTRARRKPSRLALLRTLSTVLLAFVACWGPLFLLILLD 274
Db 184 IYSLVTRSRRLTFRK ----SVSKASRSSEKSLALLKTVIIVLSVFIACWAPLFILLLD 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRDLRHALLRVCCGRHSCGRDPSG 334
Db 240 VGCKVKTCDILFYKAEYFLVAVLNSGTNPITYTLTNKEMRRAFIRIV-----SCKKCPGG 294
QY 335 SQOSASAAEASGGLRCLPPGLDGSFSGSERS-PORD 371
Db 295 -----DSTGKIKRPIIAGVEFSRSKSDNSHPKOD 324

```

```

Db 295 -----DRAQVKRPIVAGVEFSRSKSDNSHPKOD 324

RESULT 15
ID Q99N07 PRELIMINARY; PRT; 326 AA.
AC Q99N07;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE EDG1 (FRAGMENT)
GN EDG1.
OS Agouti taczanowski.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Mystricognathi; Agoutidae; Agouti.
OX NCBI_TaxID=143288;
RN [1]
RP SEQUENCE FROM N.A.
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011715; AAK01983.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36652 MW; 3AD51684BF81A889 CRC64;

Query Match 39.8%; Score 803; DB 11; Length 326;
Best Local Similarity 48.7%; Pred. No. 2e-45;
Matches 164; Conservative 62; Mismatches 93; Indels 18; Gaps 4;

QY 36 LRADAVCLAVCAFTVLENLAVLLVGRPHRPHAPMFLLSGLSLDLAGAAYANILL 95
Db 5 VKUTTVFVILICCFIILENIFVLLTWKTKRPHRPMYFIGNLALSOLLAGVAYIANLL 64
QY 96 SGLPTLKLSPALMFAREGGVVALTSVLSLALALERSLTMARRGPAPVSSRGRTLAM 155
Db 65 SGATTYKLTPAQWFLREGSMFALSASFSLAIAIERVITMLKMKLHNGSNSSRSFLL 124
QY 156 AAAAGVSLLLGLLPALGNCNCLGRDCACTVPLPYAKAYVLCVLAFCVLAIAICALYAR 215
Db 125 SACWVISLILGGLPIMGWNCISLSSCSTVPLPYHKHYILFCTVFTVTLALLAIVLYCR 184
QY 216 YCOVRANARLPARPAGTAGTSTRARRKPSRLALLRTLSTVLLAFVACWGPLFLLILLD 275
Db 185 YSLVTRSRRLTFRKNISKT-----SRSEKSLALLKTVIIVLSVFIACWAPLFILLLD 240
QY 276 ACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRDLRHALLRVCCGRHSCGRDPSG 335
Db 241 GCKVKTCDFLYKAEYFLVAVLNSGTNPITYTLTNKEMRRAFIRIM-----SCKKCPGG- 294
QY 336 SQOSASAAEASGGLRCLPPGLDGSFSGSERS-PORD 371
Db 295 -----DSTGKIKRPIIAGVEFSRSKSDNSHPKOD 324

```

Search completed: January 16, 2002, 23:43:27  
Job time: 319 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.  
OM protein - protein search, using sw model  
Run On: January 16, 2002, 23:04:03 ; Search time 58.72 Seconds  
(without alignments)  
152,526 Million cell updates/sec  
Title: US-09-842-316-2  
Perfect score: 2019  
Sequence: 1 MESGLLRPAVPSEVIVLHYN.....TGSPGAPTAARTLVSEPAAD 398  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 212252 seqs, 22503292 residues  
Total number of hits satisfying chosen parameters: 212252  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : Issued\_Patents\_AA.\*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pap.\*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pap.\*  
3: /cgn2.6/ptodata/2/1aa/6A\_COMB.pap.\*  
4: /cgn2.6/ptodata/2/1aa/6B\_COMB.pap.\*  
5: /cgn2.6/ptodata/2/1aa/PCTUS\_COMB.pap.\*  
6: /cgn2.6/ptodata/2/1aa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	853.5	42.3	383	1	US-08-196-989B-4
2	853.5	42.3	383	2	US-08-760-936-4
3	834	41.3	381	2	US-08-845-566-3
4	834	41.3	381	2	US-08-467-948A-28
5	834	41.3	381	3	US-08-852-824-18
6	834	41.3	381	3	US-08-467-947A-28
7	834	41.3	381	5	PCT-US96-10618-4
8	776	38.4	334	5	US-08-118-270-73
9	776	38.4	334	5	PCT-US93-08528-73
10	774.5	38.4	378	4	US-09-082-088-2
11	732.5	36.3	352	1	US-08-196-989B-2
12	732.5	36.3	352	2	US-08-760-936-2
13	631.5	21.3	284	3	US-08-852-824-4
14	548	21.9	509	2	US-08-845-566-1
15	518.5	25.7	364	4	US-08-763-938-2
16	508.5	25.2	351	3	US-08-861-747-2
17	505.5	25.0	393	5	PCT-US96-10618-3
18	502.5	24.9	364	5	PCT-US96-10618-2
19	499.5	24.7	351	3	US-08-789-882-2
20	445.5	22.1	354	4	US-08-325-897-2
21	443	21.9	353	3	US-08-997-803-14
22	339.5	16.8	393	3	US-08-997-803-15
23	339.5	16.8	393	3	US-08-467-948A-4
24	339.5	16.8	393	3	US-08-467-947A-4
25	304.5	15.1	345	1	US-08-118-270-70
26	304.5	15.1	345	5	PCT-US93-08528-70
27	283.5	14.0	325	3	US-08-706-281A-18

ALIGNMENTS

RESULT 1  
US-08-196-989B-4  
; Sequence 4, Application US/08196989B  
; Patent No. 5585476  
; GENERAL INFORMATION:  
; APPLICANT: MacLennan, A. John  
; TITLE OF INVENTION: Molecular Cloning and Expression of  
; TITLE OF INVENTION: G-Protein Coupled Receptors  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Saliwanchik & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/196.989B  
; FILING DATE: 15-FEB-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lloyd, Jeff  
; REGISTRATION NUMBER: 35,589  
; REFERENCE/DOCKET NUMBER: MAC-100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 383 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-196-989B-4

Query Match 42.3%; Score 853.5; DB 1; Length 383;  
Best Local Similarity 47.6%; Pred. No. 2.5e-58;  
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;  
QY 13 EVIVLHYNTKLRGARYQPGAGLRADAVVCLAVCAFTVLENLAVLVLGRHFRHPMF 72  
DB 24 DIIVRHNYNTKGL-NIGVEKDHGKLTWSVFLICCLILLENIFVLLTIWTKKFFHPMY 82  
QY 73 LLLGSLTLLDLAGAAYANILLSSPLTLKLSPALWFAREGGVFVALTASVLSLLAIALE 132

```

Db 83 YFIGNLALSDLLAGVAYTANLLSGATTYKLTTPAQWFLREGSMFVALSASFVSLAIAIE 142
QY 133 RSLTWARRGPAPVSSRGRTLAMAAAAGVSLILGLLPALGWNCIGRLDACSTVLPYAKA 192
Db 143 RYITMLKMKLHNGSNSSRFLISACWVLSILGLLPINGWNCISLSSCSTVLPYHKH 202
QY 193 YVLCFVLAFCVILAAICALYARIYCOVRANARRLPARPCTAGTTSTRARRKPRSLALLRT 252
Db 203 YILECTVFTVLLLSIVLYCRIYSLVTRSRRLTRK-----NISKASRSSEKSLALLKT 258
QY 253 LSVLLAFVACWGPLFLLLLLDVACAPCTVLLQADPGLGLAMANSLLNPIIYTLNRD 312
Db 259 VIIVLSVFIACWAPLFLLLLLDVGCRAKCDILYKAEYFLVAVLNSGTNPPIIYTLNKE 318
QY 313 LRHALLRVCCGRHSGDRPSGQSQSASAAEASGGLRRLCPGLDGSFSGSERS--PORD 371
Db 319 MRRAFIRII-----SCKKCPNG-----DSAGKFKRPIIPGMFEFSKSDNSSSHPOKD 365
QY 372 GLDTSGTSGSPG 383
Db 366 DGDNPETIMSSG 377

```

## RESULT 2

```

US-08-760-936-4
; Sequence 4, Application US/08760936
; Patent No. 5856443
; GENERAL INFORMATION:
; APPLICANT: MacLennan, A. John
; TITLE OF INVENTION: Molecular Cloning and Expression of
; G-Protein Coupled Receptors
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,936
; FILING DATE: December 6, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Doran R.
; REGISTRATION NUMBER: 38,261
; REFERENCE/DOCKET NUMBER: MAC-100C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-373-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 383 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-760-936-4

```

```

Query Match 42.3%; Score 853.5; DB 2; Length 383;
Best Local Similarity 47.6%; Pred. No. 2.5e-58;
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;
QY 13 EVIVLHNYTKRGARYQFAGLRADAVVCLAVCAFIIVLENLAVLVIGRHRPRFHAPMF 72
Db 24 DIIVRHNYTKGL-NIGVEKDGHKLTGKLVVFLICLLILENIFVLLTWKTKFHRPHY 82

```

```

QY 73 LLIGSTLSDLLAGAAAYANILLSGFLTKLSPALWFAREGGVFAVTASVLSLAIAIE 132
Db 83 YFIGNLALSDLLAGVAYTANLLSGATTYKLTTPAQWFLREGSMFVALSASFVSLAIAIE 142
QY 133 RSLTWARRGPAPVSSRGRTLAMAAAAGVSLILGLLPALGWNCIGRLDACSTVLPYAKA 192
Db 143 RYITMLKMKLHNGSNSSRFLISACWVLSILGLLPINGWNCISLSSCSTVLPYHKH 202
QY 193 YVLCFVLAFCVILAAICALYARIYCOVRANARRLPARPCTAGTTSTRARRKPRSLALLRT 252
Db 203 YILECTVFTVLLLSIVLYCRIYSLVTRSRRLTRK-----NISKASRSSEKSLALLKT 258
QY 253 LSVLLAFVACWGPLFLLLLLDVACAPCTVLLQADPGLGLAMANSLLNPIIYTLNRD 312
Db 259 VIIVLSVFIACWAPLFLLLLLDVGCRAKCDILYKAEYFLVAVLNSGTNPPIIYTLNKE 318
QY 313 LRHALLRVCCGRHSGDRPSGQSQSASAAEASGGLRRLCPGLDGSFSGSERS--PORD 371
Db 319 MRRAFIRII-----SCKKCPNG-----DSAGKFKRPIIPGMFEFSKSDNSSSHPOKD 365
QY 372 GLDTSGTSGSPG 383
Db 366 DGDNPETIMSSG 377

```

## RESULT 3

```

US-08-845-566-3
; Sequence 3, Application US/08845566
; Patent No. 5912144
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl
; TITLE OF INVENTION: EDG-1 LIKE RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,566
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0271 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 181948
US-08-845-566-3

```

```

Query Match 41.3%; Score 834; DB 2; Length 381;
Best Local Similarity 46.8%; Pred. No. 7.8e-57;

```

```

Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;
QY 13 EVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHPMF 72
DB 23 DIIVRHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHPMF 81
QY 73 LLGSLTSLDLAAGAYANILSGPLTKLSPALWFAREGGVFVALTASVLSLALALE 132
DB 82 YFIGNLALSDDLAVAYTANLLSGATTYKLTPOQWFLREGSMFVALSASVLSLALALE 141
QY 133 RSLTMARRGPAPVSSRGRTLAMAAAGVSLILGLLPALGWNCGLRDLACSTVLPYAKA 192
DB 142 RYITMKMLKHGNSNFRFLISACWISLILGGPLMGWNCISALSSTVLPYHKKH 201
QY 193 YVLCVLAFCVILAAICALYARIYCOVRANARLPARTAGTTSTRARRKPRSLALLRT 252
DB 202 YILFCTVFTLLLSIVILYCRISLVTRSRRLTKRNI-----SKASRSSENVALLKT 256
QY 253 LSVVLLAFVACGPFLLLLLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNRD 312
DB 257 VIIVLSVFIACHAPLFLLLLDVCGKVKTCIDLFRAYEFLVLAIVNSGTPNIYTLNKE 316
QY 313 LRHALLRLVCCGRHSCGRDPGSGSQSASAAEGGLRCLPPLDGSFSGSERS-PORD 371
DB 317 MRRAFRIM-----SCKKCPSG-----DSAGKFKRPIIAGMEFSRSKSDNSHPORD 363
QY 372 GLDTSGSTGSPG 383
DB 364 EGDNPETIMSSG 375

```

```

RESULT 4
US-08-467-948A-28
: Sequence 28, Application US/08467948A
: Patent No. 5998164
: GENERAL INFORMATION:
: APPLICANT: LI, YI
: APPLICANT: CAO, LIANG
: APPLICANT: NI, JIAN
: APPLICANT: GENTZ, REINER
: APPLICANT: BULT, CAROL J.
: APPLICANT: SUTTON III, GRANGER G.
: APPLICANT: ROSEN, CRAIG A.
: TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
: STREET: 1100 NEW YORK AVE., NW, SUITE 600
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY DISK
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,948A
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04079
: FILING DATE: 30-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: STEFFE, ERIC K.
: REGISTRATION NUMBER: 36,688
: REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-2500
: TELEFAX: 202-371-2540
: INFORMATION FOR SEQ ID NO: 28:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 381 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: not relevant
: MOLECULE TYPE: peptide
: US-08-467-948A-28
Query Match 41.3%; Score 834; DB 2; Length 381;
Best Local Similarity 46.8%; Pred. No. 7.8e-57;
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;
QY 13 EVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHPMF 72
DB 23 DIIVRHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHPMF 81
QY 73 LLGSLTSLDLAAGAYANILSGPLTKLSPALWFAREGGVFVALTASVLSLALALE 132
DB 82 YFIGNLALSDDLAVAYTANLLSGATTYKLTPOQWFLREGSMFVALSASVLSLALALE 141
QY 133 RSLTMARRGPAPVSSRGRTLAMAAAGVSLILGLLPALGWNCGLRDLACSTVLPYAKA 192
DB 142 RYITMKMLKHGNSNFRFLISACWISLILGGPLMGWNCISALSSTVLPYHKKH 201
QY 193 YVLCVLAFCVILAAICALYARIYCOVRANARLPARTAGTTSTRARRKPRSLALLRT 252
DB 202 YILFCTVFTLLLSIVILYCRISLVTRSRRLTKRNI-----SKASRSSENVALLKT 256
QY 253 LSVVLLAFVACGPFLLLLLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNRD 312
DB 257 VIIVLSVFIACHAPLFLLLLDVCGKVKTCIDLFRAYEFLVLAIVNSGTPNIYTLNKE 316
QY 313 LRHALLRLVCCGRHSCGRDPGSGSQSASAAEGGLRCLPPLDGSFSGSERS-PORD 371
DB 317 MRRAFRIM-----SCKKCPSG-----DSAGKFKRPIIAGMEFSRSKSDNSHPORD 363
QY 372 GLDTSGSTGSPG 383
DB 364 EGDNPETIMSSG 375

```

```

RESULT 5
US-08-852-824-18
: Sequence 18, Application US/08852824C
: Patent No. 6060272
: GENERAL INFORMATION:
: APPLICANT: LI, et al.
: TITLE OF INVENTION: Human G-Protein Coupled Receptors
: FILE REFERENCE: 1488.1220000
: CURRENT APPLICATION NUMBER: US/08/852,824C
: CURRENT FILING DATE: 1997-05-04
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 18
: LENGTH: 381
: TYPE: PRT
: ORGANISM: genomic
: US-08-852-824-18
Query Match 41.3%; Score 834; DB 3; Length 381;
Best Local Similarity 46.8%; Pred. No. 7.8e-57;
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;
QY 13 EVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHPMF 72
DB 23 DIIVRHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHPMF 81
QY 73 LLGSLTSLDLAAGAYANILSGPLTKLSPALWFAREGGVFVALTASVLSLALALE 132
DB 82 YFIGNLALSDDLAVAYTANLLSGATTYKLTPOQWFLREGSMFVALSASVLSLALALE 141

```

QY 133 RSLTWARRGPAPVSSRGRTLMAAAAGVSLILGLLPLAGWNCGLRDLACSTVPLLYAKA 192  
 Db 142 RYITMLKMLKLNNGNNFLRLLISACWVSLILGLPLIMGWNCISALSSCSTVPLLYHKH 201  
 QY 193 YLFCVLAFAVGLAAICALYARIYCOVRANARLPPARPGTAGTSTRARRKPRSLALLRT 252  
 Db 202 YILFCTTVTLLLSIVILYCRISLVRSRLTKRNI-----SKASRSSENVALLKT 256  
 QY 253 LSVLLAFVACWGLPFLLLLDVACPARCPCVLLQADPFLGLAMANSLLNPIIYTLNRD 312  
 Db 257 VIIVLSVFIACWAPFLFLLLDVGCWKTCIDLFRAYEFLVLAVLSNGTNPPIIYTLNKE 316  
 QY 313 LRHALLRVCGRHSGRDPGSGSQSASAAEASGGLRCLPGLDGSFSGSERS--PORD 371  
 Db 317 MRRAFIRM-----SCCKCPSG-----DSAGKFRPIIAGMEFRRSKSDNSHPQKD 363  
 QY 372 GLDTSSTGSGSPG 383  
 Db 364 EGDNPETIMSSG 375

RESULT 6

US-08-467-947A-28  
 ; Sequence 28, Application US/08467947A  
 ; Patent No. 6090575  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LI, YI  
 ; APPLICANT: CAO, LIANG  
 ; APPLICANT: NI, JIAN  
 ; APPLICANT: GENTZ, REINER  
 ; APPLICANT: BULT, CAROL J.  
 ; APPLICANT: SUTTON III, GRANGER G.  
 ; APPLICANT: ROSEN, CRAIG A.  
 ; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
 ; TITLE OF INVENTION: Coupled Receptor GPR1  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
 ; CITY: WASHINGTON  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: FLOPPY DISK  
 ; COMPUTER: IBM PC COMPATIBLE  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/467,947A  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/04079  
 ; FILING DATE: 30-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: STEFFE, ERIC K.  
 ; REGISTRATION NUMBER: 36,688  
 ; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 381 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: peptide  
 ; US-08-467-947A-28

Query Match 41.3%; Score 834; DB 3; Length 381;

Best Local Similarity 46.8%; Pred. No. 7.8e-57;  
 Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;  
 QY 13 EVIIVHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIYVLENLAVLLVLRHPRHAPMF 72  
 Db 23 DIIVRYHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIYVLENLAVLLVLRHPRHAPMF 81  
 QY 73 LLSGSLTDLLAGAAYAAANILLSGPLTLKSPALWFAREGGVFVALTASVLSLALIALE 132  
 Db 82 YFIGNLALSDLLAGVAYTANLLSGATTYKLTLPQWFLREGSMFVALSASFLLAIAIE 141  
 QY 133 RSLTWARRGPAPVSSRGRTLMAAAAGVSLILGLLPLAGWNCGLRDLACSTVPLLYAKA 192  
 Db 142 RYITMLKMLKLNNGNNFLRLLISACWVSLILGLPLIMGWNCISALSSCSTVPLLYHKH 201  
 QY 193 YLFCVLAFAVGLAAICALYARIYCOVRANARLPPARPGTAGTSTRARRKPRSLALLRT 252  
 Db 202 YILFCTTVTLLLSIVILYCRISLVRSRLTKRNI-----SKASRSSENVALLKT 256  
 QY 253 LSVLLAFVACWGLPFLLLLDVACPARCPCVLLQADPFLGLAMANSLLNPIIYTLNRD 312  
 Db 257 VIIVLSVFIACWAPFLFLLLDVGCWKTCIDLFRAYEFLVLAVLSNGTNPPIIYTLNKE 316  
 QY 313 LRHALLRVCGRHSGRDPGSGSQSASAAEASGGLRCLPGLDGSFSGSERS--PORD 371  
 Db 317 MRRAFIRM-----SCCKCPSG-----DSAGKFRPIIAGMEFRRSKSDNSHPQKD 363  
 QY 372 GLDTSSTGSGSPG 383  
 Db 364 EGDNPETIMSSG 375

RESULT 7

PCT-US96-10618-4  
 ; Sequence 4, Application PC/TUS9610618  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Coleman, Roger  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Au-Young, Janice  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Seilhamer, Jeffrey J.  
 ; TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 1.5  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/10618  
 ; FILING DATE: 20-JUN-1996  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/000,352  
 ; FILING DATE: 20-JUN-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/567,817  
 ; FILING DATE: 06-DEC-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Glaister, Debra J.  
 ; REGISTRATION NUMBER: 33,888  
 ; REFERENCE/DOCKET NUMBER: PF-0042 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:



INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 381 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 119130  
PCT-US96-10618-4

Query Match 41.3%; Score 834; DB 5; Length 381;  
Best Local Similarity 46.8%; Pred. No. 7.8e-57;  
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;  
QY 13 EVTLHYNTGKRGARVOPGAGLRADAVCVLAFCATVLENLAVLLVGLGHRPHAPMF 72  
DB 23 DIIVRYNTGKLNISADKENS-IKLTSVFILICCFIILENIFVLTITWTKKFRPMY 81  
QY 73 LLGSLTSLDLAAGYAANILLGSLTSLKLPALWFAREGGVFVALTASVLSLLAIALE 132  
DB 82 YFIGNLSDLLAGVAYTANLLSGATTYKLTQAQWFLREGSMFVALSASVLSLLAIAE 141  
QY 133 RSLTWARRGPAPVSSRGRTLMAAAWCVSLLGLLPGALGNCLGRDACSTVLPYAKA 192  
DB 142 RYITMKLWKLHNSNFRFLLSACWVISLILGGLPIMGNWCISALSSGCTVLPYHKH 201  
QY 193 YVLCVAVFGVLAIAICALYARIYQVKNARRLPARPAGTAGTSTRARRKPSRLALLRT 252  
DB 202 YILFCTVFTLLLSIVLYCRYISLVTRSRRLTFKNI-----SKASRSSENVALLKT 256  
QY 253 LSVLLAFVACGPFLLLLLDVACPARTCPVLLQADPFGLAMANSLLNPIIYTLNRD 312  
DB 257 VVILSVFIAWAPLFIILLDVGCKVKTCILFRAEYFLVAVLNSGTNPITYTLNKE 316  
QY 313 LRHALLRLVCCGRHSCGRDPSQSSASAEASGLRCLPGLDGSFSGSERSS-PORD 371  
DB 317 MRARFIRM-----SCKCPG-----DSAGKFRPIIAGMEFSRSKSDNSHPQKD 363  
QY 372 GLDTSSTGSGSPG 383  
DB 364 EGDNPETIMSSG 375

RESULT 8  
US-08-118-270-73  
Sequence 73, Application US/08118270  
Patent No. 5508384  
GENERAL INFORMATION:  
APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236

FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 73:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 334 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-73  
Query Match 38.4%; Score 776; DB 1; Length 334;  
Best Local Similarity 46.4%; Pred. No. 1.9e-52;  
Matches 162; Conservative 64; Mismatches 95; Indels 28; Gaps 7;  
QY 40 AVVCLAVCAFTVLENLAVLLVGLGHRPHAPMFLLGLSLTSLDLAAGYAANILLSGPL 99  
DB 3 SVYFLICCFIILENIFVLTITWTKKFRPMYFIGNLSDLLAGVAYTANLLSGAT 62  
QY 100 TLKLSPALWFAREGGVFVALTASVLSLLAIALERSLTMAR---RGPAPVSSRGRTLMAAA 156  
DB 63 TYKLPQAQWFLREGSMFVALSLSVLSLAIATERYITMLKMLHG-----SNNRFLFLIS 118  
QY 157 AAWGVSLLLGLLPGALGNCLGRDACSTVLPYAKAYLFCVL-AFVGILAAICALYARI 215  
DB 119 ACWVISLILGGLPIMGNWCISALSSGCTVLPYHKHYLFTLIVTLLLSIVLYCRI 178  
QY 216 YCOVKNARRLPARPAGTAGTSTRARRKPSRLALLRTSLSVLLAFVACGPFLLLLLDV 275  
DB 179 YSLVTRSRRLTFKNI-----SKASRSSENVALLKTIVILSVFIAWAPLFIILLD 233  
QY 276 ACPAFTCPVLLQADPFGLAMANSLLNPIIYTLNRDLRHALLRLVCCGRHSCGRDPSGS 335  
DB 234 GCKVKTCDILFRAEYFLVAVLNSGTNPITYTLNEMRRAFIRIMCC-----KCPG- 286  
QY 336 QQSASAEASGLRCLPGLDGSFSGSERSS-PORDGLDTSSTGSGSPG 383  
DB 287 -----DSAGKFRPIIAGMEFSRSKSDNSHPQKDEGDNPETIMSSG 328

RESULT 9  
PCT-US93-08528-73  
Sequence 73, Application PC/TUS9308528  
GENERAL INFORMATION:  
APPLICANT: New York University  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08528  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236



CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Lloyd, Jeff  
REGISTRATION NUMBER: 35,589  
REFERENCE/DOCKET NUMBER: MAC-100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 352 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-196-989B-2

Query Match 36.3%; Score 732.5; DB 1; Length 352;  
Best Local Similarity 45.5%; Pred. No. 4.4e-49;  
Matches 166; Conservative 48; Mismatches 122; Indels 29; Gaps 6;

QY 3 SGLLRPAPESEVIVLHNYTGTGLRGARYOPGAGLRADAVVCLAVCAFIIVLENLVLVIG 62  
Db 6 SEYLNPEKVE---HNYTKETLDQETPSRKVASAFIIL--CCAIVVENLLVLIAYA 59  
QY 63 RHRFRHAPMFLGSLTSLDLAGAAYAAANILLGSLTLPALWFAREGGVFVALTAS 122  
Db 60 RNSKFHSAMYLEFLGNLAASDLGAVAFVANTLLSGPVTLSLPLQWFAREGSAFILTAS 119  
QY 123 VLSLLAIALERSITMARRGPAPVSSRGRTLMAAAAGVSLLLGLLPALGNWCLGRDLC 182  
Db 120 VFSLLAIAIEROVAIAKVKLYGSDKSCRMMLIGASWLSILGLPLTGNWCLDHEAC 179  
QY 183 STVLPYAKAYVLCVLAFAVGLAALCAIYARICQVARNARRLPARGTAGTSTRARR 242  
Db 180 STVLPYAKHYVLCVTFISVILLALVALYVRIYFVRS-----SHADVAG----- 225  
QY 243 KPSLALLRTSLVLLAFVACWGLPLLLLDVACPARTCPVLLQADPFLGLAMANSLLN 302  
Db 226 -POTLALLTIVTLGVFIICWLPAPFISILLDSTCPVRACPVLYKAHYFFAFATLNSLLN 284  
QY 303 PIITYLTNRDLRHALLRLVCC---GRHSCGRDPSSQGSASAAEASGGLRCL-----PP 354  
Db 285 PVIYTWRSRDLRREVLRPLLCWRQKGATGRRGNPGHRLPLRSSSSSLRGLHMTSPT 344  
QY 355 GLDGS 359  
Db 345 FLEGN 349

RESULT 12  
US-08-760-936-2  
Sequence 2, Application US/08760936  
Patent No. 5856443  
GENERAL INFORMATION:  
APPLICANT: MacLennan, A. John  
TITLE OF INVENTION: Molecular Cloning and Expression of  
TITLE OF INVENTION: G-Protein Coupled Receptors  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/760,936  
FILING DATE: December 6, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Doran R.  
REGISTRATION NUMBER: 38,261  
REFERENCE/DOCKET NUMBER: MAC-100C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 352 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-760-936-2

Query Match 36.3%; Score 732.5; DB 2; Length 352;  
Best Local Similarity 45.5%; Pred. No. 4.4e-49;  
Matches 166; Conservative 48; Mismatches 122; Indels 29; Gaps 6;

QY 3 SGLLRPAPESEVIVLHNYTGTGLRGARYOPGAGLRADAVVCLAVCAFIIVLENLVLVIG 62  
Db 6 SEYLNPEKVE---HNYTKETLDQETPSRKVASAFIIL--CCAIVVENLLVLIAYA 59  
QY 63 RHRFRHAPMFLGSLTSLDLAGAAYAAANILLGSLTLPALWFAREGGVFVALTAS 122  
Db 60 RNSKFHSAMYLEFLGNLAASDLGAVAFVANTLLSGPVTLSLPLQWFAREGSAFILTAS 119  
QY 123 VLSLLAIALERSITMARRGPAPVSSRGRTLMAAAAGVSLLLGLLPALGNWCLGRDLC 182  
Db 120 VFSLLAIAIEROVAIAKVKLYGSDKSCRMMLIGASWLSILGLPLTGNWCLDHEAC 179  
QY 183 STVLPYAKAYVLCVLAFAVGLAALCAIYARICQVARNARRLPARGTAGTSTRARR 242  
Db 180 STVLPYAKHYVLCVTFISVILLALVALYVRIYFVRS-----SHADVAG----- 225  
QY 243 KPSLALLRTSLVLLAFVACWGLPLLLLDVACPARTCPVLLQADPFLGLAMANSLLN 302  
Db 226 -POTLALLTIVTLGVFIICWLPAPFISILLDSTCPVRACPVLYKAHYFFAFATLNSLLN 284  
QY 303 PIITYLTNRDLRHALLRLVCC---GRHSCGRDPSSQGSASAAEASGGLRCL-----PP 354  
Db 285 PVIYTWRSRDLRREVLRPLLCWRQKGATGRRGNPGHRLPLRSSSSSLRGLHMTSPT 344  
QY 355 GLDGS 359  
Db 345 FLEGN 349

RESULT 13  
US-08-852-824-4  
Sequence 4, Application US/08852824C  
Patent No. 6060272  
GENERAL INFORMATION:  
APPLICANT: Li, et al.  
TITLE OF INVENTION: Human G-Protein Coupled Receptors  
FILE REFERENCE: 1488.122000  
CURRENT APPLICATION NUMBER: US/08/852,824C  
CURRENT FILING DATE: 1997-05-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 384  
TYPE: PRT  
ORGANISM: genomic  
US-08-852-824-4

Query Match 31.3%; Score 631.5; DB 3; Length 384;

```
LIBRARY: TYMNOR01
CLONE: 144690
US-08-845-566-1
```

Query Match 27.18; Score 548; DB 2; Length 509;  
Best Local Similarity 46.28; Pred. No. 9.6e-35;  
Matches 121; Conservative 37; Mismatches 84; Indels 20; Gaps 5;

QY 12 SEVIVLHYNTGKLRAGRYQGAGLRADAVVCVLACAFIVLENLAVLLVGLGRHPRHPADM 71  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 22 SELIXLHYNHSRLRAGRGPDGGGLGALRGLSVASCLVLENLVLLVAITSHMSRRWV 81  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 72 FILLGSATLSDLLGAAYAAANILLSGPLTKLSPALFAREGGVFVALTASVLSLIAIAL 131  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 82 YICLVNIITSLDTGAAYLANVLSGARFTFLAPAQHFLEGILFTALAASFTSLLFTAG 141  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 132 EFSLTMARGPAPVSSRG-----RTLMAAAANGVSLLLGLLPALGWNCICLRDCACTYL 186  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 142 EFATMVR---PVAESCATKTSRVYFIGICWLALLGMLPLIGWCNICAFDRCSSEL 197  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 187 PLXAKAYLFVCVLAVFGILAIICALIARIYICOVRNARRLPARPGETAGTTSTTRARKPRS 246  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 198 PHYSKRYILFCIVFIAGVLAITMGLYGAIFRLVQAQSOKAP-RPA-----ARRKAR- 247  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 247 LALLRTLSVLLAVFACWGPLE 268  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 248 -ELLKTVMILLAFVLCWGPLE 268  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

US-08-763-958-2  
Sequence 2, Application US/08763938  
Patent No. 6140060  
GENERAL INFORMATION:  
APPLICANT: CHUN, Jerold J.M.  
APPLICANT: HECHT, Jonathan H.  
TITLE OF INVENTION: CLONED LYOPHOSPHATIDIC ACID  
TITLE OF INVENTION: RECEPTORS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray and Oram LLP  
STREET: 655 15th Street, N.W., Suite 330 - G St. Lobby  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/763,938  
FILING DATE: 12-DEC-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: JAHNS, Kristina M.  
REGISTRATION NUMBER: 41,092  
REFERENCE/DOCKET NUMBER: P8074-6018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-5000  
TELEFAX: (202) 638-4810  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 364 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-763-958-2

Search completed: January 16, 2002, 23:40:02  
Job time: 2159 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 23:35:18 ; Search time 65.2 seconds

(without alignments)  
464.992 Million cell updates/sec

Title: US-09-842-316-2

Perfect score: 2019

Sequence: 1 MESGLLRPAVPSEVIVLHYN.....TGSPGAPTAARTIVSEPAAD 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_68:\*

2: PIR1:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	853.5	42.3	383	2 I53870	Edg-1 orphan recep
2	834	41.3	381	2 A35300	G protein-coupled
3	815.5	40.4	362	3 JC7559	sphingosine 1-phos
4	774.5	38.4	378	2 JC5245	G protein-coupled
5	732.5	36.3	352	2 JC1465	probable G protein
6	506.5	25.1	364	2 JC5293	lysophosphatidic a
7	440.5	21.8	180	2 E48909	G protein-coupled
8	357	17.7	330	2 A55689	G protein-coupled
9	347	17.2	330	2 S40434	G protein-coupled
10	325.5	16.1	473	2 A33117	cannabinoid recep
11	318	15.8	362	2 I65990	G protein-coupled
12	310	15.4	472	2 S17595	cannabinoid recep
13	308	15.3	363	2 S48697	probable G protein
14	289.5	14.3	360	2 S36750	cannabinoid recep
15	282.5	14.0	372	2 I49008	melanocortin-5 rec
16	282	14.0	466	2 S36794	beta-1-adrenergic
17	281.5	13.9	325	2 I46416	melanocyte-stimula
18	278	13.8	325	2 JC5592	melanocortin 5 rec
19	275	13.6	325	2 JC2193	melanocortin recep
20	273.5	13.5	347	2 S70364	cannabinoid recep
21	273	13.5	332	2 A57055	melanocortin recep
22	272.5	13.5	464	2 S12591	beta-1-adrenergic
23	272	13.5	325	2 JN0764	melanocortin recep
24	267	13.2	480	2 I53053	beta 1 adrenergic
25	264	13.1	360	2 B46647	melanocortin recep
26	262	13.0	323	2 S36636	melanocortin recep
27	259.5	12.9	477	1 QRHUB1	beta-1-adrenergic
28	259	12.8	323	2 S43850	melanocortin 3 rec
29	259	12.8	428	2 A55044	beta-4C-adrenergic

30	257.5	12.8	440	2 JC5520	serotonin receptor
31	255.5	12.7	405	2 S65459	beta-3-adrenergic
32	252.5	12.5	374	2 I77467	serotonin receptor
33	250.5	12.4	389	2 S68422	serotonin receptor
34	247.5	12.3	320	2 S18444	G protein-coupled
35	247.5	12.3	386	2 S18637	serotonin receptor
36	246.5	12.2	386	2 S54153	serotonin receptor
37	246	12.2	317	2 T12055	melanocyte stimula
38	244.5	12.1	386	2 A42688	serotonin receptor
39	244.5	12.1	390	2 JN0268	serotonin receptor
40	243.5	12.1	418	2 G02953	beta-3-adrenergic
41	243.5	12.1	560	2 A38731	alpha-1A adrenergic
42	243	12.0	390	2 S58126	serotonin receptor
43	243	12.0	437	2 I57942	5-hydroxytryptamin
44	243	12.0	445	2 T42203	serotonin receptor
45	240	11.9	517	2 A45121	alpha-1B adrenergic

## ALIGNMENTS

RESULT 1

I53870

Edg-1 orphan receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999

C:Accession: I53870

R:Lado, D.C.; Browe, C.S.; Gaskin, A.A.; Borden, J.M.; MacLennan, A.J.

A:Title: Cloning of the rat edg-1 immediate-early gene: expression pattern suggests d

A:Reference number: I53870; MUID:95047498

A:Accession: I53870

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-383 <RES>

A:Cross-references: EMBL:U0303; NID:g595396; PIDN:AAA83418.1; PID:g595397

C:Superfamily: G protein-coupled receptor edg-1

Query Match 42.3%; Score 853.5; DB 2; Length 383;  
Best Local Similarity 47.6%; Pred. No. 1.2e-58;  
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;

QY 13 EVIVLHYNTGKLRGARYQPGAGLRADAVCLVCAFTVLENLAVLLVGRHPRFHPMF 72

DB 24 DIIVRHYNTGKLRGARYQPGAGLRADAVCLVCAFTVLENLAVLLVGRHPRFHPMF 82

QY 73 LLGSLTLDLLAGAAAYANILLGSLTGLKLPALWFAREGGVFVALTSVLSLALALE 132

DB 83 YFIGNALSDLLAGVAYTANLLSGATTYKLTPAQWFLREGSMFVALSVFSLALALE 142

QY 133 RSLTWARRGPAPVSRGRTRLAMAAANGVSLLLGLLPALGNCLGRDLACSTVLPYAKA 192

DB 143 RYITMLKMKLHNGSSRSRSLISACWVLSLILGLPIMGWNCISSLSGCTVLPYHKH 202

QY 193 VYLFVCLAVFGVTLAICALYARIYCOVRANARRLPARGTAGTSTTRARKPRSLALLRT 252

DB 203 YILFTTTVFTLLLSLIVILYRIYSLVTRSRRLTFRK----NISKASRSSEKSLALLKT 258

QY 253 LSVLLAFVACWGPFLFLLLDVACPARTCPVLQADPPFLGLAMANSLLNPIIYTTNRD 312

DB 259 VIIVLSVFIACWAPFLFLLLDVGCARTCDILYKAEVFLVAVLNSGTNPVIYTLNKE 318

QY 313 LRHALLRLVCCGRISCGDPSGQSSASAAASGGLRCLPPLGLDGSFGSERSSS--PQD 371

DB 319 MRAPRIII-----SCCKCPNG-----DSAGKFRPIIPGMEFSRSKSNSSHPQKD 365

QY 372 GLDTSGTSGSPG 383

DB 366 DGDNPETIMSSG 377

RESULT 2

A35300  
G protein-coupled receptor edg-1 - human  
C:Species: Homo sapiens (man)  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 24-Sep-1999  
C:Accession: A35300  
R:HLA, T.; Maciag, T.  
J. Biol. Chem. 265, 9308-9313, 1990  
A:Title: An abundant transcript induced in differentiating human endothelial cells encoded by a novel human gene  
A:Reference number: A35300; MUID:90264425  
A:Accession: A35300  
A:Molecule type: mRNA  
A:Residues: 1-381 <HLA>  
A:Cross-references: GB:M31210; NID:g181948; PIDN:AAA52336.1; PID:g181949  
C:Genetics:  
A:Gene: GDB:ECGF1  
A:Cross-references: GDB:127754; OMIM:131222  
A:Map position: 22q13-22q13  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein

Query Match 41.3%; Score 834; DB 2; Length 381;  
Best Local Similarity 46.8%; Pred. No. 4e-57;  
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;

QY 13 EVIVLYNTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHAPMF 72  
DB 23 DIIVRYNTGKLRNTSADKENS-IKLTVSVFILCCFLENIENIFVLTITWTKKFRPMY 81  
QY 73 LLGSLTSDLLAGAAAYANILLSPGLTKLSPALWFAREGGVFVALTASVLSLLALE 132  
DB 82 YFIGNLSDLLAGVAYTANLLSGATTYKLTPTQWFLREGSMFVALSASVFSLLAIE 141  
QY 133 RSLTMARRGPAPVSSRGRTLAMAAWGVSLGLLPGALWNCGLRDLACSTVLPYAKA 192  
DB 142 RYITMLKMLHNGSNFRLFLISACWVLSLGLPLMGWNCISALSSTVLPYHKH 201  
QY 193 YVLCVLAFFVGLAALCALYARIYCOVRANARLPARPCTAGTTSTRARKPRSLALLRT 252  
DB 202 YILFCTVTLTLLSVILYCHYSLVTRSRRLTFKNI-----SKASRSSENVALKLT 256  
QY 253 LSVLAFVACWGPFLILLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYTTNRD 312  
DB 257 VIIVSVFTACWAPFLILLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYTTNKE 316  
QY 313 LRHALLRLVCCGRHSGRDPGSGSQSASAAEAGGLRCLPPGLDGSFSGSERSS-PQRD 371  
DB 317 MRRAPRIM-----SCCKCPSG-----DSAGKFRPIIAGMEFSRKSNDSSHPQKD 363  
QY 372 GLDTSGTSGSPG 383  
DB 364 EGDNPETIMSSG 375

RESULT 3  
JC7559  
sphingosine 1-phosphate receptor - zebra fish  
N:Alternate names: endothelial differentiation gene 1 receptor  
C:Species: Brachydanio rerio (zebra fish)  
C:Date: 31-Mar-2001 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: JC7559  
R:Im, D.S.; Ungar, A.R.; Lynch, K.R.  
Biochem. Biophys. Res. Commun. 279, 139-143, 2000  
A:Title: Characterization of a zebrafish (Danio rerio) sphingosine 1-phosphate receptor  
A:Reference number: JC7559; MUID: 20563813  
A:Accession: JC7559  
A:Molecule type: DNA  
A:Residues: 1-362 <IMA>  
A:Cross-references: GB:A3321294  
C:Comment: This receptor, a G protein-coupled receptor, mediates sphingosine 1-phosphate or also functions in the developing central nervous system in cell proliferation, apoptosis  
C:Genetics:  
A:Gene: edg1

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 40.4%; Score 815.5; DB 3; Length 362;  
Best Local Similarity 47.1%; Pred. No. 1e-55;  
Matches 172; Conservative 59; Mismatches 107; Indels 27; Gaps 7;

QY 11 VSEIVLYNTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHAP 70  
DB 1 MODLLARHYNFTGKRVKHKDP--GLKADSVVFIIVCCFIENLAVLLVLRHPRHAP 58  
QY 71 MFLLLSGTSLDLAGAAAYANILLSPGLTKLSPALWFAREGGVFVALTASVLSLLAIA 130  
DB 59 MYFICNLALSLLAGVYVYTNILLSGANTYKLTPTQWFFREGSMFVALAASVFLSLLAIA 118  
QY 131 LERSLTMAARRGPAPVSSRGRT--LAWAAAAGVSLLLGLLPGALWNCGLRDLACSTVLP 187  
DB 119 IERHUTMK---MKLHNGKTCRVFMLISTVWFIAAILGLLPGVWNCIDISINNCSVLP 175  
QY 188 LYAKAYVLCVLAFFVGLAALCALYARIYCOVRANARLPARPCTAGTTSTRARKPRSL 247  
DB 176 LYHKAYILFCTVFSVILMAIVLYARIYALVTRSRKLVPRKVANGSGNKSSEK--SM 233  
QY 248 ALLRTLSVLLAFVACWGPFLILLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYT 307  
DB 234 ALLKTVIIVLSCTACWAPFLILLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYT 293  
QY 308 LTNDRLLRALLVCCGRHSGRDPGSGSQSASAAEAGGLRCLPPGLDGSFSGSERSS 367  
DB 294 LLSNEMRRAPFKMLNCG-----VCVQPSGKFSRPI-MGAEFSTKSDNSS 337  
QY 368 -PQRD 371  
DB 338 HENKD 342

RESULT 4

JC5245  
G protein-coupled receptor - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: JC5245  
R:Yamaguchi, F.; Tokuda, M.; Hatase, O.; Brenner, S.  
Biochem. Biophys. Res. Commun. 227, 608-614, 1996  
A:Title: Molecular cloning of the novel human G protein-coupled receptor (GPCR) gene  
A:Reference number: JC5245; MUID:97032811  
A:Accession: JC5245  
A:Molecule type: DNA  
A:Residues: 1-378 <YAN>  
A:Cross-references: EMBL:X83864; NID:g1770395; PIDN:CAA58744.1; PID:g1770396  
C:Comment: This protein transduces hormone and neurotransmitter signals into intracellular  
C:Genetics:  
A:Gene: edg-3  
A:Map position: 9q22.1-q22.2  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor

Query Match 38.4%; Score 774.5; DB 2; Length 378;  
Best Local Similarity 44.7%; Pred. No. 1.5e-52;  
Matches 156; Conservative 55; Mismatches 109; Indels 29; Gaps 4;

QY 6 LRPAVSEIVLYNTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPR 65  
DB 9 LQVRGNETLREHYQYVGLKAGRLKEASEGSTLTITVFLVICSFIENLAVLLVLRHPR 68  
QY 66 RFHAPMFLLSGLTSLDLAGAAAYANILLSPGLTKLSPALWFAREGGVFVALTASVLS 125  
DB 69 KFNRMVFFIGNLALCDLLAGIAYKVNILSGKTKFSLSPVWFLREGSMFVALGASTCS 128  
QY 126 LLATALERSLTMAARRGPAPVSSRGRTLAMAAWGVSLLLGLLPGALWNCGLRDLACSTV 185  
DB 129 LLATAIERHLLTMKRPYDANKRHRVFLIIGMCMWLTFTLGLPILGWNCLNLPDCSTI 188



Qy 355 GLDGS 359  
1:1:

Matches	97;	Conservative	27;	Mismatches	53;	Indels	15;	Gaps	2;
QY	66	RHPAPMFLGLSLTSDLLAGAAANILLSPGLTKLSPALMFAREGGVFVALTSVLS	125						
Db	2	KPHSAYFLFGLNLAASDLLAGVAFVANTLLSGHVTLSLTVPQFAREVSNAFTLSASVFS	61						
QY	126	LLAIALERSLTHWARGPAPVSSRGRTLMAAAANGVSLLLGLLPALGWNCGLRDACTSV	185						
Db	62	LLAIAIEROVALAKYKLYGSDKSCRMMLLIGASWLSLILGLSLILGWNCNKLEACSTV	121						
QY	186	LPLXAKAYVLCVAFVGLIAAICALYARYICVQRANARLPGCTAGTSTTRARRKPR	245						
Db	122	LPLXAKRYVLCVTTFSVILLAIIVALYRIYFVVRS-----SHADVAG-----	166						
QY	246	SLALLRTLSWVL	257						
Db	167	TLALLKTTIVL	178						

RESULT 8

A53689

G protein-coupled receptor 3 - human

N:Alternate names: G protein-coupled receptor GPR3; orphan G-protein-coupled receptor AC

C:Species: Homo sapiens (man)

C>Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 19-May-2000

C:Accession: A55689; S58521; C55733

C:Rismaa, T.P.; Klefer, J.; Liu, M.L.; Baker, E.; Sutherland, G.R.; Shine, J.

Genomics 24, 391-394, 1994

A:Title: Isolation and chromosomal localization of a novel human G-protein-coupled receptor

A:Reference number: A55689; MUID:95213036

A:Accession: A55689

A:Status: preliminary

A:Molecule type: DNA; mRNA

A:Residues: 1-330 <IIS>

R:Cross-references: GB:L32830; GB:L32831; NID:g602311; PIDN:AAA73560.1; PID:g602312

R:Eggensickx, D.; Denet, J.F.; Labbe, O.; Hayashi, Y.; Refetoff, S.; Vassart, G.; Parmant

Biochem. J. 309, 837-843; 1995

A:Title: Molecular cloning of an orphan G-protein-coupled receptor that constitutively a

A:Reference number: S58521; MUID:95366960

A:Accession: S58521

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-330 <EGG>

A:Cross-references: GB:X83956; NID:g1061125; PIDN:CAA58787.1; PID:g1061126

R:Marchese, A.; Docherty, J.M.; Nguyen, T.; Heiber, M.; Cheng, R.; Heng, H.H.Q.; Tsui, I.

Genomics 23, 609-618, 1994

A:Title: Cloning of human genes encoding novel G protein-coupled receptors.

A:Reference number: A55733; MUID:95154831

A:Accession: C55733

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-292 <MAR>

A:Cross-references: GB:U13668; NID:g577416; PIDN:AAA64594.1; PID:g577417

C:Genetics:

A:Gene: GDB:GPR3

A:Cross-references: GDB:371695; OMIM:600241

A:Map position: lp36.1-1p34.3

C:Superfamily: melanocortin receptor

C:Keywords: G protein-coupled receptor; transmembrane protein

```

Query Match      17.7%; Score 357; DB 2; Length 330;
Best Local Similarity 35.5%; Pred. No. 2.3e-20;
Matches 107; Conservative 40; Mismatches 108; Indels 46; Gaps 12;

QY   39 DAVGLVCATVLENTAIVLVLRGRPRFHAPMFLLLGSLTLDLAG---AYAAANTILL 95
      | :|::: | | :: | | | | | | | | | | | | | | | | | | | |
DB   44 DVYLGIS-GTLVSCENALVAIIIVGTPTAFRAPMFLVLGVSLAVADLLAGLGLVLFHPAAVFC 102
      | :|::: | | :: | | | | | | | | | | | | | | | | | | | |
QY   96 SGPLTKLSPALWTAAREGGVFVALTSVLSLIAIALERSLTWARR----GPAPVSSRGRT 151
      | :|::: | | :: | | | | | | | | | | | | | | | | | | | |
db   103 IGSAEWSL-----VLVGVLMAFTASIGSLAITVDRLYSLNALTYSETTVT---RT 153

```

```

QY      152 LAMAAAAAGVSLLLGLLPALGWNCLGRDLACSTVPLYAKAYVLFCVLAFVGILAAICAL 211
        ||| : ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      154 YVMALVWGGAUGLGLLPVALWNLCDGLTTCGVVVYPL- SKNHLVLAIAFAFMVFGIMQL 212
        ||| : ||||| | | | | | | | | | | | | | | | | | | | | | | | |

QY      212 YA---RIYC-----QVRANARLPAFPGTAGTTSTRARRKPRSLALLRTLTVLLAFVAVCW 264
        ||| : ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      213 YAOICRIVCRHAQQTALQRHLLPA-----SHYVATRGK-----IATLAVLVGAFAPAC 260
        ||| : ||||| | | | | | | | | | | | | | | | | | | | | | | | |

QY      265 GPFLFLLLLDVA--CPARTCFVLQADPFLGLAMANSLLNPITYTLTNRDLRHALRLVC 322
        ||| : ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      261 LPPTYVCLLGDAHSPLYTYITLLP-----ATYNSMINPIYAFRNQDVOKVLNAVCC 313
        ||| : ||||| | | | | | | | | | | | | | | | | | | | | | | | |

QY      323 C 323
        |
Db      314 C 314

RESULT          9
S40454
G protein-coupled receptor GPCR21 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1994 [sequence_revision 10-Nov-1995 #text_change 20-Jun-2000]
C;Accession: S40454
R;Saeki, Y.; Ueno, S.; Mizuno, R.; Nishimura, T.; Fujimura, H.; Nagai, Y.;
FEBS Lett. 336, 317-322, 1993
A;Title: Molecular cloning of a novel putative G protein-coupled receptor
A;Reference number: S40454; MUID:S4085630

```

## RESULT

G protein-coupled receptor GPCR21 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 19-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 20-Jun-2  
 C:Accession: S40454  
 P:Saeki, Y.; Ueno, S.; Mizuno, R.; Nishimura, T.; Fujimura, H.; Nagai, Y  
 FEBS Lett. 336, 317-322, 1993  
 A:Title: Molecular cloning of a novel putative G protein-coupled receptor  
 A:Reference number: S40454; MUID:94085630  
 A:Accession: S40454  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-330 <SAE>  
 A:Cross-references: GB:D21062; NID:9455487; PIDN:BAA04641.1; PID:g455488  
 C:Superfamily: melanocortin receptor

```

Query Match      17.2%; Score 347; DB 2; Length 330;
Best Local Similarity 34.3%; Pred. No.1.4e-119;
Matches 103; Conservative 43; Mismatches 110; Indels 44; Gaps 11;

QY      39  DAVVCLAVCAFTVLGNLAVLLVGLGRHFAPMPELLGSLTLDLLAGAA----YAANIL 94
      . | | : : | | : : | | | | | | | | | | | | | | | | | | | |
Db       44  DVVLGIS-GLTVSCENALVVAIICTPAFRAPMELLVGLSLAVADLGLGLVLFHFAOFC 102

QY      95  LSGPLTLKLSPALWFARGGVFVALTSVLSLLAIALERSLTWARR-----GRAPVSSRR 150
      : | | : | | : | | | | | | | | | | | | | | | | | | | |
Db      103  IG-----SPMSLMLVGLVLAAMAFATASIGSLLAITVDRLSYLNALTYSETTVT--R 152

QY      151  TLAMAAAANGVSLLLGLPALGWCNGLRDLACSVLPYAKAVLCVCLAFVGLLAACA 210
      . | | : | | | | | | | | | | | | | | | | | | | | | | |
Db      153  TYVMALVWVGALGLGPLVLAWNCRDGLTTCGVVYPL-SKNHLVLVAIAFEMVGINLQ 211

QY      211  LVA---RVC-----QVRANARLPARGCTAGTSTARRAPRSALLRLSVLLAFVAC 263
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      212  LYAQICRIVCRHAQIALQRHLHP-----SHYVATRG-----IATLAVVLGAFAAC 259

QY      264  WGPLFLLLLDVACPARTCPVLQADPFLGLMANSLNLPYITLTNRDLRALLRLVCC 323

Db      260  WLPFTYVLCIGDADSPRLTYLT-----LIPATVNSMINPVIKAFNODVQVLAICC 314

```

RESULT 10

```

RESULT 10
A33117
C:cannabinoid receptor CB1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 05-Nov-1999
C:Accession: A33117; C55879
R:Natsuda, I.A.; Lolait, S.J.; Brownstein, M.J.; Young, A.C.; Bonner, T.I.
Nature 346, 561-564, 1990
A:title: Structure of a cannabinoid receptor and functional expression of t
A:Reference number: A33117; MUID:90332039
A:Accession: A33117
A:Molecule type: mRNA

```

A:Residues: 1-473 <MAT>  
A:Cross-references: GB:X55812; NID:g1552375; PIDN:CAA39332.1; PID:g57249  
R:Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.; Ca  
J. Biol. Chem. 270, 3726-3731, 1995  
A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from a  
A:Reference number: A55879; MUID:95181329  
A:Accession: C55879  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-107 <SH>  
C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 16.1%; Score 325.5; DB 2; Length 473;  
Best Local Similarity 30.1%; Pred. No. 8.9e-18;  
Matches 91; Conservative 59; Mismatches 117; Indels 35; Gaps 9;

QY 38 ADVVCLVCAPIVLENLAVLLVGLGRPHFA-----PMFLLGSLTSLDLAGAAVAA 91  
DB 119 AIAVLSLTGTTVLENLLVLCV-----LHSLRLCRPSYHFGSLAVADLLGSLVIFY 173  
QY 92 NILLGSLTLKSLPALFARGGVFVALTSVLSLAIALERSLTWARRGPAPVS----- 146  
DB 174 SFYDFHVFHRKDSPPNVFLKGGVTASFTASVSLFTADRYVISHR-----PLAYKRIV 229  
QY 147 SGRGTLMAAAAGVSLLLGLPALGNCLGRDLACSTVPLAKAYVLCVLAFFVGLA 206  
DB 230 TRPKAVAFCLMTATVIAVPLGLGNCKLQSVCSDFPLIDETILYFMWIGTVSVLL 289  
QY 207 AICALYARIYCOVRANARRLPARPGTAGT-----TS-----TRARRKPSRLALLRTSLV 255  
DB 290 FIYAVMYLTKAHSHAVRMIO-RTQKSIITSEDKGVQVTRPDQARMDIRLAKTLVL 348  
QY 256 VLLAFVACGPLELILLDLVACPARTCPVLLQADPLFG-LAMANSLLNPIIYTLNDRLR 314  
DB 349 ILVLLICNGPLLAIVYDF--GKMKLITVFACSMCLLNSTVNPPIYALRSKDLR 406  
QY 315 HA 316  
DB 407 HA 408

RESULT 11  
165990  
G protein-coupled receptor 6 - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999  
C:Accession: I65990  
R:Heiber, M.; Docherty, J.M.; Shah, G.; Nguyen, T.; Cheng, H.H.Q.; Marchese, A.  
DNA Cell Biol. 14, 25-35, 1995  
A:Title: Isolation of three novel human genes encoding G protein-coupled receptors.  
A:Reference number: I53033; MUID:95134333  
A:Accession: I65990  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:I36150; NID:g598156; PIDN:AAA63181.1; PID:g598157  
C:Genetics:  
A:Gene: GDB:GPR6  
A:Cross-references: GDB:371713; OMIM:600553  
A:Map position: 6q21-6q21  
C:Superfamily: melanocortin receptor

Query Match 15.8%; Score 319; DB 2; Length 362;  
Best Local Similarity 24.1%; Pred. No. 2.6e-17;  
Matches 104; Conservative 38; Mismatches 125; Indels 38; Gaps 9;

QY 32 PGAGLRA----DAVVCVCAPIVLENLAVLLVGLGRPHFA-----PMFLLGSLTSLDLAGAA 87  
DB 65 PGLLLPVPNDVLLCVS-GTVIAGENALVALLIATSPALRTPMEVGLSVLADTLAAGC 123  
QY 88 AYAAVLLSGPLTKLSPALWFARREGGVFALTSVLSLALALERSLTWARRGPAPVS 147

A:Residues: 1-473 <MAT>  
A:Cross-references: GB:X55812; NID:g1552375; PIDN:CAA39332.1; PID:g57249  
R:Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.; Ca  
J. Biol. Chem. 270, 3726-3731, 1995  
A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from a  
A:Reference number: A55879; MUID:95181329  
A:Accession: C55879  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-107 <SH>  
C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 15.4%; Score 310; DB 2; Length 472;  
Best Local Similarity 27.8%; Pred. No. 1.4e-16;  
Matches 96; Conservative 65; Mismatches 142; Indels 42; Gaps 11;

QY 38 ADVVCLVCAPIVLENLAVLLVGLGRPHFA-----PMFLLGSLTSLDLAGAAVAA 91  
DB 119 AIAVLSLTGTTVLENLLVLCV-----LHSLRLCRPSYHFGSLAVADLLGSLVIFY 173  
QY 92 NILLGSLTLKSLPALFARGGVFVALTSVLSLAIALERSLTWARRGPAPVS----- 146  
DB 174 SFYDFHVFHRKDSPPNVFLKGGVTASFTASVSLFTADRYVISHR-----PLAYKRIV 229  
QY 147 SGRGTLMAAAAGVSLLLGLPALGNCLGRDLACSTVPLAKAYVLCVLAFFVGLA 206  
DB 230 TRPKAVAFCLMTATVIAVPLGLGNCKLQSVCSDFPLIDETILYFMWIGTVSVLL 289  
QY 207 AICALYARIYCOVRANARRLPARPGTAGT-----TS-----TRARRKPSRLALLRTSLV 255  
DB 290 FIYAVMYLTKAHSHAVRMIO-RTQKSIITSEDKGVQVTRPDQARMDIRLAKTLVL 348  
QY 256 VLLAFVACGPLELILLDLVACPARTCPVLLQADPLFG-LAMANSLLNPIIYTLNDRLR 314  
DB 349 ILVLLICNGPLLAIVYDF--GKMKLITVFACSMCLLNSTVNPPIYALRSKDLR 406  
QY 315 HA 316  
DB 407 HA 408

RESULT 12  
S17595  
cannabinoid receptor CB1 - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jul-2000  
C:Accession: S17595; S13668; A55879  
R:Gerard, C.M.; Mollereau, C.; Vassart, G.; Parmentier, M.  
Biochem. J. 279, 129-134, 1991  
A:Title: Molecular cloning of a human cannabinoid receptor which is also expressed in  
A:Reference number: S17595; MUID:92028798  
A:Accession: S17595  
A:Molecule type: mRNA  
A:Residues: 1-472 <GER1>  
A:Cross-references: EMBL:X54937; NID:g29914; PIDN:CAA38699.1; PID:g29915  
R:Gerard, C.; Mollereau, C.; Vassart, G.; Parmentier, M.  
Nucleic Acids Res. 18, 7142, 1990  
A:Title: Nucleotide sequence of a human cannabinoid receptor cDNA.  
A:Reference number: S13668; MUID:91088303  
A:Accession: S13668  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-472 <GER2>  
A:Cross-references: EMBL:X54937; NID:g29914; PIDN:CAA38699.1; PID:g29915  
R:Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.;  
J. Biol. Chem. 270, 3726-3731, 1995  
A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from  
A:Reference number: A55879; MUID:95181329  
A:Accession: A55879  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-124 <SHI>  
A:Cross-references: GB:X81120  
C:Genetics:  
A:Gene: GDB:CNRI; CNR  
A:Cross-references: GDB:127354; OMIM:114610  
A:Map position: 6q14-6q15  
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
F:117-142/Domain: transmembrane #status predicted <TM1>  
F:155-175/Domain: transmembrane #status predicted <TM2>  
F:188-212/Domain: transmembrane #status predicted <TM3>  
F:233-256/Domain: transmembrane #status predicted <TM4>  
F:275-299/Domain: transmembrane #status predicted <TM5>  
F:345-365/Domain: transmembrane #status predicted <TM6>  
F:378-399/Domain: transmembrane #status predicted <TM7>  
F:77.83/Binding site: carbohydrate (Asn) #status predicted

Db 118 AIAVLSLTGTTFTVLENNLLVLCVI-----LHRSLECRPSYHFIGSLAVADLLGSVIFVY 172  
 QY 92 NILSGPLTKLSPALWFAREGGVFVALFASVLSLLAIALERSLTMARRGPAPVS----- 146  
 Db 173 SFIDFVHFHRKSRNVFLFKLGVTAFSTFASVGSFLTAIDRYSIHR-----PLAYKRIV 228  
 QY 147 SRGRTLMAAANGVSLILGLLPAWGNCGLGRDACSTVPLPYAKAYVLCVLAFAVGILA 206  
 Db 229 TRPKAVAFCLMTIIVIAVLAFLPGWNCCKLQSVCSDFPHIDITYLWFWIGVTSVLL 288  
 QY 207 AICALYARIYCOVRANARLPPARPAGT-----TS-----TRARRKPRSLALLRTLSV 255  
 Db 289 FIVAYMYILWKAHSHAVRMIOQ-PTOKSIIITHTSEDKGVQVTRPDQARMDIRLAKTIVL 347  
 QY 256 VLLAFVACWGPLLLLLLDVACPARTCPVLLQADPFLG-LAMANSLLNPIIYTTNDRDLR 314  
 Db 348 ILVVLICWGLLAIMYDVF--GRMKLIKTVFAFCMCLLNSTVNPPIIYALRSKDLR 405  
 QY 315 HALLRLV--CCGR-----HSCGRDPSSGQSSASAAEAGGLRRCL 352  
 Db 406 HAFRSPFSCGTAQPLDNSMGSDCLHKKHANNAASVHRAAESCI 450  
 RESULT 13  
 S48697  
 probable G protein-coupled receptor protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 24-Sep-1999  
 C:Accession: S48697  
 R:Song, Z.H.; Young III, W.S.; Brownstein, M.J.; Bonner, T.I.  
 FEBS Lett. 351, 375-379, 1994  
 A:Title: Molecular cloning of a novel candidate G protein-coupled receptor from rat brain  
 A:Reference number: S48697; MUID:94364507  
 A:Accession: S48697  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-363 <SON>  
 A:Cross-references: EMBL:U012006; NID:9551333; PIDN:AAA21870.1; PID:9551334  
 C:Superfamily: melanocortin receptor  
 C:Keywords: G protein-coupled receptor

Query Match 15.3%; Score 308; DB 2; Length 363;  
 Best Local Similarity 32.9%; Pred. No. 1.5e-16;  
 Matches 100; Conservative 40; Mismatches 126; Indels 38; Gaps 9;  
 QY 34 AGRADAV---VCLAVCAFIYL-ENLAVLLVGRHPRHAPMELLGLSLTLSDLAGAA 88  
 Db 66 SGLLLSAVNPWDVLCVSGTVIAGENALVVALIATSTPAIRTPMEVVLVGLSATADLLAGCG 125  
 QY 89 YAAIILLGPTLKLSPALWFAREGGVFVALFASVLSLLAIALERSLTMARRGPAPVSSR 148  
 Db 126 LILHFVFOYVWP---SETVSLIMVGLVAFSAASVSSLLAITVDYLSLYN---ALTYYS 179  
 QY 149 GRTL-----AMAAANGVSLILGLLPAWGNCGLGRDACSTVPLPYAKAYVLCVLAFAVGI 204  
 Db 180 RRTLLGVHLLATWTVSLGLGLLPLVGNCLADRSQVVRPL-TRSHVALLSTFFEVV 238  
 QY 205 LAAICALYARIYCOVRANARLPPARPAGTGTSTRARRKPRSLALLRTLSVLLAFVACW 264  
 Db 239 FGIMLHLVYRICOVVRHAHQIALQOHCPLAPHLAATRKG-----VCTLAVVLGTGASW 293  
 QY 265 GPLFLLLLLDVACPARTCPVLLQADPFLG-----AMANSLLNPIIYTTNDRHALL 318  
 Db 294 LRF-----AIYCVGSGQEDPAIYTYATLLPATYNSMINPIIYAFRQEIQRALW 342  
 QY 319 RLVC 322  
 Db 343 LLFC 346

RESULT 14  
 S36750

cannabinoid receptor CB2 - human  
 N:Alternate names: cannabinoid receptor, peripheral  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Jan-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999  
 C:Accession: S36750  
 R:Munro, S.; Thomas, K.L.; Abu-Shaar, M.  
 Nature 365, 61-65, 1993  
 A:Title: Molecular characterization of a peripheral receptor for cannabinoids.  
 A:Reference number: S36750; MUID:93368659  
 A:Accession: S36750  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-360 <MUN>  
 A:Cross-references: EMBL:X74328  
 C:Superfamily: melanocortin receptor  
 Query Match 14.3%; Score 289.5; DB 2; Length 360;  
 Best Local Similarity 29.3%; Pred. No. 4.1e-15;  
 Matches 95; Conservative 50; Mismatches 150; Indels 29; Gaps 11;  
 QY 38 ADAVVCVAVCAFIIVLENLAVL-LVLGRHPRHAPMELLGLSLTLSDLAGAAVAAAILLS 96  
 Db 35 AVAVLCTLLGLLSALENVAVLYLSSHOLRRKPSYLFIGSLAGADEFLASVVFACSFVNF 94  
 QY 97 GPLTKLSPALWFAREGGVFVALFASVLSLLAIALERSLTMARRGPA--PVSSRGRTLAM 154  
 Db 95 HVFHVDSKAVFLKIGSVTMTFTASVGSLLLTATDRYLCL-RYPPSYKALLTRGRGLVT 153  
 QY 155 AAAANGVSLILGLLPAWGNCGLGRDACSTVPLPYAKAYVLCVLAFAVGLAATCALYAR 214  
 Db 154 LGIMVLSALVSYPLMGWTCPPR--PCSELFPLIPNDYLLSNLLFTAFIFSGIITYYGH 211  
 QY 215 YCOVRANARLPPAR-----PGTAGTSTRARRKPRSLALLRTLSVLLAFVACWGPLFL 269  
 Db 212 VLMKAHOHVASLSGHQDQVFGMA-----RMRLDVRLAKTLGLVLAVLLICWCFVLA 263  
 QY 270 LLLLDVACPARTCPVLLQADPFLG-LAMANSLLNPIIYTTNDRDLR---HALLRLVCCG 324  
 Db 264 LMAHSLA--TTLSQVKKAFACFCSMLCLINSWNPVIYALRSGEIRSSAHHLAHKKCV 321  
 QY 325 RHSCGRD--PSGQSSASAAEASG 346  
 Db 322 R-GLGSEAKEAPRSSVTETADG 344  
 RESULT 15  
 I49008  
 melanocortin-5 receptor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999  
 C:Accession: I49008; A54245; J02244  
 R:Fathi, Z.; Iben, L.G.; Parker, E.M.  
 Neurochem. Res. 20, 107-113, 1995  
 A:Title: Cloning, expression, and tissue distribution of a fifth melanocortin receptor  
 A:Reference number: I49008; MUID:95258173  
 A:Accession: I49008  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-372 <RES>  
 A:Cross-references: EMBL:U08354; NID:9522165; PIDN:AAA76585.1; PID:9522166  
 R:Labbe, O.; Desarnaud, F.; Eggerickx, D.; Vassart, G.; Parmentier, M.  
 Biochemistry 33, 4543-4549, 1994  
 A:Title: Molecular cloning of a mouse melanocortin 5 receptor gene widely expressed 1  
 A:Reference number: A54245; MUID:94213827  
 A:Accession: A54245  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 48-120, F', 122-372 <LAB>  
 A:Cross-references: GB:X76295; NID:9498973; PIDN:CAA53943.1; PID:9498974  
 A:Experimental source: clone HMP01B  
 A:Note: sequence extracted from NCBI backbone (NCBI:145988)  
 R:Gantz, I.; Shimoto, Y.; Konda, Y.; Dickinson, C.J.; Yamada, T.

Biochem. Biophys. Res. Commun. 200, 1214-1220, 1994  
A:Title: Molecular cloning, expression, and characterization of a fifth melanocortin receptor  
A:Reference number: JC2244; MUID:94241974  
A:Accession: JC2244

A:Molecule type: DNA  
A:Residues: 48-372 <GAN>  
A:Cross-references: GB:122527; NID:9468377; PIDN:AAA21337.1; PID:9468378  
C:Comment: This protein responds to melanocortins with an increase in intracellular cyclic AMP  
C:Superfamily: melanocortin receptor  
C:Keywords: receptor; transmembrane protein

F:84-109/Domain: transmembrane #status predicted <TM1>  
F:121-144/Domain: transmembrane #status predicted <TM2>  
F:164-185/Domain: transmembrane #status predicted <TM3>  
F:206-226/Domain: transmembrane #status predicted <TM4>  
F:232-257/Domain: transmembrane #status predicted <TM5>  
F:287-310/Domain: transmembrane #status predicted <TM6>  
F:324-344/Domain: transmembrane #status predicted <TM7>

Query Match 14.0%; Score 282.5; DB 2; Length 372;  
Best Local Similarity 26.2%; Pred. No. 1.5e-14;  
Matches 85; Conservative 70; Mismatches 102; Indels 67; Gaps 14;  
QY 33 GAGLRADAVVC-----LAVCAFI-----VLENALVLLGRHPRFHAPMFLLLGSLTSLD 82  
DB 70 GSNVKNKSLACEENGIAVEVFLTGLVSLLENTLVIGAIKKNKLNHSPMYFYVGS LAVAD 129  
QY 83 LLAGAAYAANILLGSPPLTK-LSPALMFAR-----EGGVEVALTASVLSLALALERSL 135  
DB 130 MLVSMNSAWETVTIYLLNNKHLVADTFVPHIDNVFDSMICISVYVASC SLAIADVRI 189  
QY 136 TM-----ARRGPAPVSSRGRTLAWAAAAGVSLLLGLLPALGN-CIGRLDACS 183  
DB 190 TIFYALRYHHIMTARRS-----GVIIACIWTFCI-----SCG 221  
QY 184 TVLPYAKA-YVLPFCVLA-FVGILAAICAYRIYCOVRANARLPARPCTAGTTSTRAR 241  
DB 222 IVFIYVESKYVICLISMEFTMLFWWSLYTHMFLARNHVKRTAASP---RYNSVRQR 278  
QY 242 RKPRSLALLRTLSVLLAFVACNGPLFLLLLLDVACPART-CPVLIQ-ADPFGLAMANS 299  
DB 279 TSMKGAI---TLTMLLGIFVCHWSPFHLILMISCPQNVYSCFMSYFNMYLILIMONS 335  
QY 300 LLNPIYITNDRDLRHALLRLVCC 323  
DB 336 VIDPLIYALRSQEMRRRTFREIVCC 359

Search completed: January 16, 2002, 23:41:21  
Job time: 363 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 16, 2002, 23:38:53 ; Search time 64.41 Seconds  
(without alignments)  
226.558 Million cell updates/sec

Title: US-09-842-316-2

Perfect score: 2019

Sequence: 1 MESGLRPAPVSEVIVLHYN.....TGSPCAPTAARTLVSEPAAD 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	865.5	42.9	382	1	EDG1_MOUSE
2	853.5	42.3	383	1	EDG1_RAT
3	834	41.3	381	1	EDG1_HUMAN
4	774.5	38.4	378	1	EDG3_HUMAN
5	732.5	36.3	352	1	H218_RAT
6	518.5	25.7	364	1	EDG2_MOUSE
7	506.5	25.1	364	1	EDG2_HUMAN
8	505.5	25.0	364	1	EDG2_BOVIN
9	505.5	25.0	393	1	EDG2_SHEEP
10	440.5	21.8	180	1	EDG1_MOUSE
11	357	17.7	330	1	GPR3_HUMAN
12	352.5	17.5	334	1	GPRC_HUMAN
13	350	17.3	334	1	GPRC_MOUSE
14	347	17.2	330	1	GPR3_MOUSE
15	347	17.2	468	1	CB1A_FUGRU
16	346.5	17.2	334	1	GPRC_RAT
17	342	16.9	470	1	CB1B_FUGRU
18	325.5	16.1	473	1	CB1R_MOUSE
19	325.5	16.1	473	1	CB1R_RAT
20	323.5	16.0	472	1	CB1R_FELCA
21	322.5	16.0	473	1	CB1R_POEGU
22	318.5	15.8	473	1	CB1R_TARGR
23	318.5	15.8	362	1	GPR6_HUMAN
24	310	15.4	472	1	CB1R_HUMAN
25	308	15.3	363	1	GPR6_RAT
26	291.5	14.4	360	1	CB2R_HUMAN
27	289.5	14.3	294	1	MC5R_PIG
28	287.5	14.2	325	1	MC5R_BOVIN
29	283.5	14.0	325	1	MC5R_MOUSE
30	283	14.0	325	1	MC5R_HUMAN
31	282	14.0	466	1	BIAR_MOUSE
32	281.5	13.9	325	1	MC5R_SHEEP
33	.281	13.9	466	1	BIAR_RAT

34	277	13.7	325	1	MC5R_PANTR	O9tt23 pan troglod
35	276	13.7	473	1	BIAR_CANFA	P79148 canis famil
36	275	13.6	325	1	MC5R_RAT	P35345 rattus norv
37	275	13.6	332	1	MC4R_PIG	O97504 sus scrofa
38	273.5	13.5	347	1	CB2R_MOUSE	P47936 mus musculu
39	273	13.5	332	1	MC4R_BOVIN	O99138 bos taurus
40	273	13.5	467	1	BIAR_BOVIN	O9tt96 bos taurus
41	272.5	13.5	360	1	CB2R_RAT	O9qzn9 rattus norv
42	272.5	13.5	474	1	BIAR_FELCA	O9tst6 felis silve
43	272	13.5	332	1	MC4R_HUMAN	P32245 homo sapien
44	269	13.3	360	1	MC3R_HUMAN	P41968 homo sapien
45	267	13.2	480	1	BIAR_MACMU	P47899 macaca mula

## ALIGNMENTS

RESULT	1
EDG1_MOUSE	
ID	EDG1_MOUSE
AC	O08530; STANDARD; PRT; 382 AA.
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUL-1998 (Rel. 36, Last annotation update)
DE	PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.
GN	EDG1.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_Taxid:10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-BALB/C; TISSUE=Liver;
RX	MEDLINE=97369927; PubMed=9226368;
RA	Liu C.H., Hia T.;
RT	"The mouse gene for the Inducible G-protein-coupled receptor edg-1.";
RL	Genomics 43:15-24(1997).
CC	-!- FUNCTION: THIS INDUCIBLE EPITHELIAL CELL G-PROTEIN-COUPLED RECEPTOR MAY BE INVOLVED IN THE PROCESSES THAT REGULATE THE DIFFERENTIATION OF ENDOTHELIAL CELLS (BY SIMILARITY).
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES WITH HIGHEST LEVELS IN BRAIN, SPLEEN AND HEART. LOWER LEVELS FOUND IN LUNG, PLACENTA, MUSCLE, LIVER, UTERUS AND KIDNEY. VERY LOW LEVELS IN TESTIS AND THYMUS.
CC	-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	EMBL; U40811; AAC53294.1; -
DR	GCRDb; GCR_1661; -
DR	MGD; MGI:1096355; Edg1.
DR	InterPro; IPR000276; GPCR_Rhodpsn.
DR	Pfam; PF00601; 7tm_1; 1.
DR	PRINTS; PRG0237; GPCRHDOPSN.
DR	PRINTS; PRG0642; EDGIORPHANR.
DR	PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR	PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;
FT	Phosphorylation; Lipoprotein; Palmitate.
FT	DOMAIN 1..46
FT	TRANSMEM 47..71
FT	DOMAIN 72..78

FT TRANSMEM 160 185 4 (POTENTIAL).  
 FT DOMAIN 186 201 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 202 222 5 (POTENTIAL).  
 FT DOMAIN 223 256 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 257 278 6 (POTENTIAL).  
 FT DOMAIN 279 294 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 295 315 7 (POTENTIAL).  
 FT DOMAIN 316 382 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 30 30 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT LIPID 328 328 PALMITATE (BY SIMILARITY).  
 FT MOD\_RES 353 353 PHOSPHORYLATION (POTENTIAL).  
 SQ SEQUENCE 382 AA; 42571 MW; A004893491C8957 CRC64;

Query Match 42.9%; Score 865.5; DB 1; Length 382;  
 Best Local Similarity 48.5%; Pred. No. 4.9e-50;  
 Matches 181; Conservative 66; Mismatches 105; Indels 21; Gaps 6;

QY 13 EVIVLHNYTKGLR-GARYOPGAGLRADAVCLVCAFTVLENLAVLLVGRHPRHAPM 71  
 DB 23 DIIVRHNYTKGLNIG--EKDHGKILTSVVFILICCFILLENIFVLLTIWTKKFRHPM 80

QY 72 FLGLSITLSDLAGAAYANILLSGPLTKLSPALWFAREGGVFVALTSVLSLALAL 131  
 DB 81 YFIGNLALSDLAGAVYATNALLSGATYKLTPAQWFLREGSMFVALSASVLSLALAI 140

QY 132 ERSITMARPPAPVSSRGRTLMAAANGVSLLLGLPALGNCLGRDAGSTVLPYAK 191  
 DB 141 ERYITMLKMLKHNSSRSFLLISACWISLILGLPSMGWNCISLSSCSTVLPYHK 200

QY 192 AYVLCVLAFLVGLIAALCALYARIYQVRANARLPARPAGTGTSTRARKPRSLALLR 251  
 DB 201 HYILFCTVTLLLSLAILCYRIYSLVTRSRRLTRFK-----NISKGRSEKSLALLK 256

QY 252 TLSVLLAFVACGPFLLILLDVPARTCPVLLQADPFLGLAMANSLLPIIYTLNR 311  
 DB 257 TVIIVSVFACWAPFILLLDVGCARKTCDILYKAELVFLVAVLNSGTPPIIYTLNK 316

QY 312 DLHALLRLVCCGRHSGRDPSCSQSQAASAGLRCPLPPGLDGFSGSERSS-POR 370  
 DB 317 EMRAFAIRIV-----SCCKCPNG-----DSAGKFRPIIPGMEFSRSDNSHPQK 363

QY 371 DGLDTSSTGSPG 383  
 DB 364 DGDNPETIMSSG 376

RESULT 2  
 EDGL\_RAT  
 ID EDGL\_RAT STANDARD; PRT; 383 AA.  
 AC P48303;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.  
 GN EDG1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RC MEDLINE=95047498; PubMed=7959012;  
 RA Lado D.C., Browe C.S., Gaskin A.A., Borden J.M., MacLennan A.J.;  
 RT "Cloning of the rat edg-1 immediate-early gene: expression pattern  
 suggests diverse functions.";  
 RL Gene 149:331-336(1994).  
 CC -1- FUNCTION: IT IS POSSIBLE THAT IT ACTS AS A RECEPTOR FOR A TROPHIC  
 CC -1- FACTOR AND IS INVOLVED IN THE SURVIVAL OF BRAIN CELLS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT EMBRYONIC DAY 15. AT  
 CC POSTNATAL DAY 14 DETECTED IN SKIN, SPLEEN, LIVER, KIDNEY, HEART,

TESTICLE, LUNG AND BRAIN. AT ADULTHOOD IS MOST ABUNDANT IN BRAIN.  
 -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: U10303; AAA83418.1; --  
 CC GCRDB; GCR\_1453;  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PRINTS; PR00642; EDGLORPHANR.  
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR\_F1\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR\_F2\_1; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.  
 FT DOMAIN 1 47 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 48 72 1 (POTENTIAL).  
 FT DOMAIN 73 79 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 80 108 2 (POTENTIAL).  
 FT DOMAIN 109 122 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 123 141 3 (POTENTIAL).  
 FT DOMAIN 142 160 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 161 186 4 (POTENTIAL).  
 FT DOMAIN 187 202 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 203 223 5 (POTENTIAL).  
 FT DOMAIN 224 257 6 (POTENTIAL).  
 FT TRANSMEM 258 279 7 (POTENTIAL).  
 FT DOMAIN 280 295 7 (POTENTIAL).  
 FT TRANSMEM 296 316 7 (POTENTIAL).  
 FT DOMAIN 317 383 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 31 31 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT LIPID 329 329 PALMITATE (BY SIMILARITY).  
 FT MOD\_RES 354 354 PHOSPHORYLATION (POTENTIAL).  
 SQ SEQUENCE 383 AA; 42745 MW; 090BA6AE09DB4F3 CRC64;

Query Match 42.3%; Score 853.5; DB 1; Length 383;  
 Best Local Similarity 47.6%; Pred. No. 3e-49;  
 Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;

QY 13 EVIVLHNYTKGLR-GARYOPGAGLRADAVCLVCAFTVLENLAVLLVGRHPRHAPM 72  
 DB 24 DIIVRHNYTKGL-NIGVEKDHGKILTSVVFILICCFILLENIFVLLTIWTKKFRHPM 82

QY 73 LILGLSITLSDLAGAAYANILLSGPLTKLSPALWFAREGGVFVALTSVLSLALAL 132  
 DB 83 YFIGNLALSDLAGAVYATNALLSGATYKLTPAQWFLREGSMFVALSASVLSLALAI 142

QY 133 RSLTMRARPPAPVSSRGRTLMAAANGVSLLLGLPALGNCLGRDAGSTVLPYAKA 192  
 DB 143 RYITMLKMLKHNSSRSFLLISACWISLILGLPIMGWNCISLSSCSTVLPYHKH 202

QY 193 VYLCVLAFLVGLIAALCALYARIYQVRANARLPARPAGTGTSTRARKPRSLALLR 252  
 DB 203 YILFCTVTLLLSLAILCYRIYSLVTRSRRLTRFK-----NISKGRSEKSLALLK 258

QY 253 LSVLLAFVACGPFLLILLDVPARTCPVLLQADPFLGLAMANSLLPIIYTLNRD 312  
 DB 259 VIIVLSVFIACWAPFILLLDVGCARKTCDILYKAELVFLVAVLNSGTPPIIYTLNKE 318

QY 313 LEHALLRLVCCGRHSGRDPSCSQSQAASAGLRCPLPPGLDGFSGSERSS-POR 371  
 DB 319 MFAIRIRII-----SCCKCPNG-----DSAGKFRPIIPGMEFSRSDNSHPQK 365

QY 372 GLDTSSTGSPG 383  
 DB 366 DGDNPETIMSSG 377

RESULT 3  
EDG1\_HUMAN STANDARD; PRT; 381 AA.  
AC P21453;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.  
GN EDG1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Umbilical vein endothelial cells;  
RX MEDLINE=90264425; PubMed=2160972;  
RA Hla T.; MacIag T.;  
RT "An abundant transcript induced in differentiating human endothelial  
RT cells encodes a polypeptide with structural similarities to  
RT G-protein-coupled receptors."  
RL J. Biol. Chem. 265:9308-9313(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=98072391; PubMed=9409733;  
RA An S.; Bleu T.; Huang W.; Hallmark O.G.; Coughlin S.R.; Goetzl E.J.;  
RT "Identification of cDNAs encoding two G protein-coupled receptors for  
RT lysophingolipids".  
RL FEBS Lett. 417:279-282(1997).  
CC -1- FUNCTION: THIS INDUCIBLE EPITHELIAL CELL G-PROTEIN-COUPLED  
CC RECEPTOR MAY BE INVOLVED IN THE PROCESSES THAT REGULATE THE  
CC DIFFERENTIATION OF ENDOTHELIAL CELLS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS, AND TO A LESSER EXTENT, IN  
CC VASCULAR SMOOTH MUSCLE CELLS, FIBROBLASTS, MELANOCYTES, AND CELLS  
CC OF EPITHELIOID ORIGIN.  
CC -1- INDUCTION: BY THE TUMOR PROMOTER PHORBOL 12-MYRISTATE 13-ACETATE  
CC (PME) IN THE PRESENCE OF CYCLOHEXIMIDE.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M31210; AA52336.1; -  
DR EMBL; AF022137; AAC51905.1; -  
DR PIR; A35300; A35300.  
DR GCRDB; GCR\_0067; -  
DR MIM; 601974; -  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm.1.1  
DR PRINTS; PR00237; GPCRHOOPS.  
DR PRINTS; PR00642; EDGIORPHANR.  
DR PROSITE; PS00237; G-PROTEIN\_RECP\_FL1; 1.  
DR PROSITE; PS00262; G-PROTEIN\_RECP\_FL2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 46  
FT TRANSMEM 47 71  
FT DOMAIN 72 78  
FT TRANSMEM 79 107  
FT DOMAIN 108 121  
FT TRANSMEM 122 140  
FT TRANSMEM 141 159  
FT TRANSMEM 160 185  
FT TRANSMEM 186 201  
FT DOMAIN

FT TRANSMEM 202 222 5 (POTENTIAL).  
FT DOMAIN 223 255 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 256 277 6 (POTENTIAL).  
FT DOMAIN 278 293 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 294 314 7 (POTENTIAL).  
FT DOMAIN 315 381 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 30 30 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 36 36 N-LINKED (GLCNAC...) (POTENTIAL).  
FT LIPID 327 327 PALMITATE (BY SIMILARITY).  
FT MOD\_RES 352 352 PHOSPHORYLATION (POTENTIAL).  
SQ SEQUENCE 381 AA; 42695 MW; D706DA4C74C2E093 CRC64;  
Query Match 41.3%; Score 834; DB 1; Length 381;  
Best Local Similarity 46.8%; Pred. No. 5, 6e-48;  
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;  
QY 13 EVIVLYNTGKLGARYQPGAGLRADAVVCAVIVLENLAVLLVGRHPRHAPMF 72  
Db 23 DIIVRYNTGKLNISADKENS-IKLTSVVFLICCFIILENIFVLLTWTKKFRPMY 81  
QY 73 LLGSLTSLDLAAGAYANILLGSLTLLKLSPALFAWFAEGGVFALTASVLSLATALE 132  
Db 82 YFIGNLALSDDLAVAYTANLLSGATTYKLTPAQMFLEGGSMFVALSASFSLATAIE 141  
QY 133 RSLTMARRGPAPVSSRGRTLMAAAAANGVSLLLGLPALGNCLGRDLACSTVLPYAKA 192  
Db 142 RYITLKMKLHNGSNFRFLILISACWISLILGLPIMGNCISALSSCSTVLPYHKK 201  
QY 193 YVLCVLAFAVGIILAIYARIYQVNRANRLPARPTAGTTSTRARRKPRSLALRT 252  
Db 202 YILFCTVFTLLLSIVLYCRYSLVTRSRRLTFRKNI-----SKASRSSENVALKLT 256  
QY 253 LSVLLAFAVACGFLFLLLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNRD 312  
Db 257 VIIVLVFATACWAPFILLLLDVGCKVKTCDLFRAYFLVLAIVNSGTNPDIYTLNKE 316  
QY 313 LRHALRLVCCGRHSGRDPSPGSGQSASAAEASGGLRRLCPGLDPSGSGSERSS-PORD 371  
Db 317 MRRAFIRIN-----SCKKCPSG-----DSAGKFKRPIIAGMEFSRSKSDNSHPKOD 363  
QY 372 GLDTSGSTGSPG 383  
Db 364 EGDNPETIMSSG 375

RESULT 4  
EDG3\_HUMAN STANDARD; PRT; 378 AA.  
AC Q99500;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYOSOPHINGLIPID RECEPTOR (EDG-3).  
GN EDG3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=97032811; PubMed=8878560;  
RA Yamaguchi F.; Tokuda M.; Hatase O.; Brenner S.;  
RT "Molecular cloning of the novel human G protein-coupled receptor  
RT (GPCR) gene mapped on chromosome 9".  
RL Biochem. Biophys. Res. Commun. 227:608-614(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=98072391; PubMed=9409733;  
RA An S.; Bleu T.; Huang W.; Hallmark O.G.; Coughlin S.R.; Goetzl E.J.;  
RT "Identification of cDNAs encoding two G protein-coupled receptors for







Db 79 RFHPIYLLMANLAAADFFAGLAYFYLMTGPNTRRLTVSTWLLRGLDITSLTASVAN 138  
 QY 126 LLAIALERSUTWARCPAPVSSRGHTLMAAAANGVSLLLGLLPALGWNCLGRDACSTV 185  
 Db 139 LLAIAIERHITVFMOLHTRMSNRVVVVIWTHMAIVMGAIPSVGVNCCIDHCNSNM 198  
 QY 186 LPLAKAYVLCVAFVGLAALCALYARIYCOVRANARLPARPAGTGTSTRARRKPR 245  
 Db 199 APLYSDSLVAFWAIFNLVTFVVMVLYAHIFGYVQRTMRSR-----SSGPRNRDT 252  
 QY 246 SLALLRTLSVLLAFVACWGLFLLLLLDVACPARTCPVLLQADPFGLAMANSLLNP 305  
 Db 253 MMSLLKTVVILGAFIVCTPGLVLLLDVCCP--QCDVLAKEKFFLLAEFNSANNP 310  
 QY 306 YLTNRDLRHALLRLVCCGRSCGRDPS--GSOOSASA 341  
 Db 311 YSYRDKEMSAFTRQILCCORNENGTGSDRSASS 347

## RESULT 7

## EDG2\_HUMAN

ID EDG2\_HUMAN STANDARD; PRT; 364 AA.

AC Q92633; O00656; P78351;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE LYOPHOSPHATIDIC ACID RECEPTOR (EDG-2).

GN EDG2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97224397; PubMed=9070858;

RA An S., Dickens M.A., Bleu T., Hallmark O.H., Goetzi E.J.;

RT "Molecular cloning of the human Edg2 protein and its identification

as a functional cellular receptor for lysophosphatidic acid.;"

RL Biochem. Biophys. Res. Commun. 231:619-622(1997).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=97224241; PubMed=9069262;

RA Moolenaar W.H., Kranenburg O., Postma F.R., Zondag G.C.M.;

RT "Lysophosphatidic acid: G-protein signalling and cellular responses.;"

RL Curr. Opin. Cell Biol. 9:168-173(1997).

CC -1- FUNCTION: RECEPTOR FOR LYOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF

DIVERSE CELLULAR ACTIVITIES.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; U08011; AAC51139.1; -

DR EMBL; Y09479; CAAY70686.1; -

DR EMBL; Y09479; CAAY70687.1; -

DR EMBL; U78192; AAC00530.1; -

DR GCRDB; GCR\_1342; -

DR GCRDB; GCR\_1897; -

DR GCRDB; GCR\_1898; -

DR MIM; 602282; -

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm\_1; 1.

DR PRINTS; PR01148; LPARECEPTOR.

DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.

DR PROSITE; PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.

DR G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50  
 FT TRANSSEM 51 75  
 FT DOMAIN 76 82  
 FT TRANSSEM 83 111  
 FT DOMAIN 112 125  
 FT TRANSSEM 126 144  
 FT DOMAIN 145 163  
 FT TRANSSEM 164 189  
 FT DOMAIN 190 205  
 FT TRANSSEM 206 226  
 FT DOMAIN 227 258  
 FT TRANSSEM 259 280  
 FT DOMAIN 281 294  
 FT TRANSSEM 295 315  
 FT DOMAIN 316 364  
 FT LIPID 327 327  
 FT CARBOHYD 35 35  
 FT CARBOHYD 35 35  
 FT CONFLICT 340 340  
 SQ SEQUENCE 364 AA; 41139 MW; 59E772FD00DFE74 CRC64;

Query Match 25.1%; Score 506.5; DB 1; Length 364;

Best Local Similarity 32.9%; Pred. No. 1.4e-26;

Matches 112; Conservative 73; Mismatches 130; Indels 25; Gaps 5;

QY 12 SEVIVLHYNTGKRGARYQP-----GAGLRADAVYVCAVFIYLVLENLAVLLVGRHP 65  
 Db 27 NESIAFFYNSRGKHLATENTVSKVLGLG-----ITVCIFIMLANLLVMVAIYVNR 78  
 QY 66 EFHAPFLILGSLTDLLAGAAYANILLGSLTLKLSPALWFAREGGVFVALTASVLS 125  
 Db 79 EFHPIYLLMANLAAADFFAGLAYFYLMETGPNTRRLTVSTWLLRQGLDITSLTASVAN 138  
 QY 126 LLAIALERSUTWARCPAPVSSRGHTLMAAAANGVSLLLGLLPALGWNCLGRDACSTV 185  
 Db 139 LLAIAIERHITVFMOLHTRMSNRVVVVIWTHMAIVMGAIPSVGVNCCIDHCNSNM 198  
 QY 186 LPLAKAYVLCVAFVGLAALCALYARIYCOVRANARLPARPAGTGTSTRARRKPR 245  
 Db 199 APLYSDSLVAFWAIFNLVTFVVMVLYAHIFGYVQRTMRSR-----SSGPRNRDT 252  
 QY 246 SLALLRTLSVLLAFVACWGLFLLLLLDVACPARTCPVLLQADPFGLAMANSLLNP 305  
 Db 253 MMSLLKTVVILGAFIVCTPGLVLLLDVCCP--QCDVLAKEKFFLLAEFNSANNP 310  
 QY 306 YLTNRDLRHALLRLVCCGRSCGRDPSGSOOSASAAEAS 345  
 Db 311 YSYRDKEMSAFTRQILCCOR---SENPTGPTESSDRSASS 347

## RESULT 8

## EDG2\_BOVIN

ID EDG2\_BOVIN STANDARD; PRT; 364 AA.

AC Q28031;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE LYOPHOSPHATIDIC ACID RECEPTOR (EDG-2) (RECI.3).

GN EDG2.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=Brain;

RX MEDLINE=97165887; PubMed=9013780;

RA Macrae A.D., Premont R.T., Jaber M., Petersen A.S., Lefkowitz R.J.;

RT "Cloning, characterization, and chromosomal localization of recl.3, a

member of the G-protein-coupled receptor family highly expressed in

RT brain.;"

Brain Res. Mol. Brain Res. 42:245-254(1996).

-1- FUNCTION: RECEPTOR FOR LYSOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF DIVERSE CELLULAR ACTIVITIES.

-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U48236; AAC48695.1; -

CC GCRDB; GCR\_1218; -

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm1; 1.

DR PRINTS; PRO1148; LPARECEPTOR.

DR PROSITE; PS00237; G-PROTEIN\_RECF\_F1\_1; 1.

DR PROSITE; PS00262; G-PROTEIN\_RECF\_F1\_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 51 75 1 (POTENTIAL).

FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 83 111 2 (POTENTIAL).

FT DOMAIN 112 125 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 126 144 3 (POTENTIAL).

FT DOMAIN 145 163 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 164 189 4 (POTENTIAL).

FT DOMAIN 190 205 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 206 226 5 (POTENTIAL).

FT DOMAIN 227 258 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 259 280 6 (POTENTIAL).

FT DOMAIN 281 294 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 295 315 7 (POTENTIAL).

FT DOMAIN 316 364 CYTOPLASMIC (POTENTIAL).

FT LPID 327 327 PALMITATE (BY SIMILARITY).

FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 364 AA; 41070 MW; DE94675BA407B2F6 CRC64;

Query Match 25.0%; Score 505.5; DB 1; Length 364;

Best Local Similarity 33.5%; Pred. No. 1.6e-26;

Matches 113; Conservative 70; Mismatches 131; Indels 23; Gaps 5;

QY 12 SEVIVLHYNTGKLRGARYQP-----GAGLRADAVVCLAVCAFTVLENLAVLLVLRGHP 65

DB 27 NESIAFFYNSRGKYLATENWTVTKLVMLG-----ITVCIFIMLANLLVWVAIYVNR 78

QY 66 RFIAPFELLGSLTSLDLAAGAAAYANILLGSLPLKLSPALFWAREGVFVALTSVLS 125

DB 79 RFHPIYILMANLAADFFLAGLYFYMFGTNTGTRRLTSTVTLRGLDITSLTVSVAN 138

QY 126 LLAIALERSLTMARPGAPVSSRGRTLMAAAAGVSLGLLGLPALGNCLGRDLDCSTV 185

DB 139 LLAIAIERHTVTRQOLHARNSRRVVIVITWVAIVMGAIPIVSGNCTCDIENCSNM 198

QY 186 LPIYAKAYLVFCVLAFLAICALYARIYCOVRANRRLPAREGTAGTTSTRARRKPR 245

DB 199 APLYSDSYLVFWAIFNLVTFVWVLYAHIFGYVRQTRMRSH-----SSGPRNRDT 252

QY 246 SLALLRTLVLVLAFCVAGPLFLLLLDVACPARTCPVLQADPFLGLANSLNLPIL 305

DB 253 MWSELKTVLVGLAFCITCPGLVLLLDVCCP--QCDVLAWEKFLLLAEFNSANPIL 310

QY 306 YTLNRLRLHALLRLVCCGRHSCGRDPS--GSQOQSASA 341

DB 311 YSYRDKMSATFRQILCCQRSENSTSPTEGSDRSASS 347

RESULT 9

EDG2\_SHEEP

ID EDG2\_SHEEP STANDARD; PRT; 393 AA.

AC P46628;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE LYSOPHOSPHATIDIC ACID RECEPTOR (EDG-2).

GN EDG2 OR EDG-2.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI\_TaxID:9940;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Folitary pars tuberalis;

RX MEDLINE=96430916; PubMed=883398;

RA Masana M.I., Brown R.C., Pu H., Gurney M.E., Dubocovich M.L.;

RT "Cloning and characterization of a new member of the G-protein coupled receptor EDG family."

RL Recept. Channels 3:255-262(1995).

CC -1- FUNCTION: RECEPTOR FOR LYSOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF DIVERSE CELLULAR ACTIVITIES.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U18405; AAB52368.1; -

DR GCRDB; GCR\_1271; -

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm1; 1.

DR PRINTS; PRO1148; LPARECEPTOR.

DR PROSITE; PS00237; G-PROTEIN\_RECF\_F1\_1; 1.

DR PROSITE; PS00262; G-PROTEIN\_RECF\_F1\_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 51 75 1 (POTENTIAL).

FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 83 111 2 (POTENTIAL).

FT DOMAIN 112 125 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 126 144 3 (POTENTIAL).

FT DOMAIN 145 163 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 164 189 4 (POTENTIAL).

FT DOMAIN 190 205 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 206 226 5 (POTENTIAL).

FT DOMAIN 227 258 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 259 280 6 (POTENTIAL).

FT DOMAIN 281 294 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 295 315 7 (POTENTIAL).

FT DOMAIN 316 364 CYTOPLASMIC (POTENTIAL).

FT LPID 327 327 PALMITATE (BY SIMILARITY).

FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 393 AA; 44433 MW; 356E961153C345FC CRC64;

Query Match 25.0%; Score 505.5; DB 1; Length 393;

Best Local Similarity 33.5%; Pred. No. 1.7e-26;

Matches 113; Conservative 70; Mismatches 131; Indels 23; Gaps 5;

QY 12 SEVIVLHYNTGKLRGARYQP-----GAGLRADAVVCLAVCAFTVLENLAVLLVLRGHP 65

DB 27 NESIAFFYNSRGKYLATENWTVTKLVMLG-----ITVCIFIMLANLLVWVAIYVNR 78

```

QY 66 RHAPMELLGSLTSLDLAGAAYANILLSGPLTLKLSPALWFAREGGVFFVALTASVLS 125
DB 79 REFFIITLMANLAADFAAGLAFYLMFTGNTRTLATVSTWLLROGLDITVTVASVAN 138

QY 126 LLAIALERSLTMRGAPVSSRGRTLAMAAAAGVSLILGLLPALGNCNCLGRDACSIV 185
DB 139 LLAIAIERHITVERMOLHTRMSNRVVVVIWIAVNGAIPSGVNCICDIENCSNM 198

QY 186 LPLAKAYVLCVLAFCVLAICALYARIYCOVRANARLARPCTAGTTSTRARRKPR 245
DB 199 APLYSDSYLVFWAIFNLVTFVVMVLYAHIFGVORTMMSRH-----SSGPRNRDT 252

QY 246 SLALRTLSVLLAFVACWGLFLLLLDVACPARTCPVLLQADPFLGLAMNSLNPIT 305
DB 253 WMSLKTVVIVLGAFCIWFPGVLVLLLDVCCP--QCDVLAYEKFLLAEFNSAMNPIT 310

QY 306 YLTNRDRHALLRLVCGRHSGRDP--GSOOSASA 341
DB 311 YSRDKEMSATFROILACORSENTSGTSGDRSASS 347

RESULT 10
EDGL_MOUSE STANDARD; PRT; 180 AA.
AC P52592;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1 LIKE (FRAGMENT).
GN GPCR13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=94116980; PubMed=8288218;
RA Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,
RA Copeland N.G., Jenkins N.A.
RT "Identification, chromosomal location, and genome organization of
RT mammalian G-protein-coupled receptors."
RL Genomics 18:175-184(1993).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20334; AAA16846.1; -
DR GCRDB; GCR_0811; -
DR MGD; MGI:95559; Gpcr13
DR InterPro; IPR000226; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN_RECF_F1.1; PARTIAL.
DR PROSITE; PS0262; G-PROTEIN_RECF_F1.2; 1.
KW G-protein coupled receptor; Transmembrane.
FT NON_TER 1 1
FT DOMAIN <1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 6 34 2 (POTENTIAL).
FT DOMAIN 35 48 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 49 67 3 (POTENTIAL).
FT DOMAIN 68 86 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 87 112 4 (POTENTIAL).
FT DOMAIN 113 128 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 129 149 5 (POTENTIAL).
FT DOMAIN 150 172 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 173 >180 6 (POTENTIAL).

```

```

FT NON_TER 180 180
SQ SEQUENCE 180 AA; 19364 MW; 91D99CF8769C4CBD CRC64;

Query Match 21.8%; Score 440.5; DB 1; Length 180;
Best Local Similarity 50.5%; Pred. No. 1.5e-22;
Matches 97; Conservative 27; Mismatches 53; Indels 15; Gaps 2;

QY 66 RHAPMELLGSLTSLDLAGAAYANILLSGPLTLKLSPALWFAREGGVFFVALTASVLS 125
DB 2 KHSAMTYLFGNLAASDLLAGVAFVANTLLSGHVTLSLTPVQVFAREVSFAITLSASVFS 61

QY 126 LLAIALERSLTMRGAPVSSRGRTLAMAAAAGVSLILGLLPALGNCNCLGRDACSIV 185
DB 62 LLAIAIERHITVERMOLHTRMSNRVVVVIWIAVNGAIPSGVNCICDIENCSNM 121

QY 186 LPLAKAYVLCVLAFCVLAICALYARIYCOVRANARLARPCTAGTTSTRARRKPR 245
DB 122 LPLAKRYVLCVVTIESVILLAIVALYVRIYFVRS-----SHADVAG-----PQ 166

QY 246 SLALRTLSVVL 257
DB 167 TLALKTVITVL 178

RESULT 11
GPR3_HUMAN STANDARD; PRT; 330 AA.
AC P46089;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE G PROTEIN-COUPLED RECEPTOR GPR3 (ACCA ORPHAN RECEPTOR).
GN GPR3 OR ACCA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96015070; PubMed=8530049;
RA Iismaa T.P., Kiefer J., Liu M.L., Baker E., Sutherland G.R.,
RA Shine J.;
RT "Isolation and chromosomal localization of a novel human G-protein-
RT coupled receptor (GPR3) expressed predominantly in the central
RT nervous system."
RL Genomics 24:391-394(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213036; PubMed=7698767;
RA Iismaa T.P., Kiefer J., Liu M.L., Baker E., Sutherland G.R.,
RA Shine J.;
RT "Isolation and chromosomal localization of a novel human G-protein-
RT coupled receptor (GPR3) expressed predominantly in the central
RT nervous system."
RL Genomics 24:391-394(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95366960; PubMed=7639700;
RA Eggerickx D., Deneef J.F., Labbe O., Hayashi Y., Refetoff S.,
RA Vassart G., Parmentier M., Libert F.;
RT "Molecular cloning of an orphan G-protein-coupled receptor that
RT constitutively activates adenylylate cyclase."
RL Biochem. J. 309:837-843(1995).
RN [4]
RP SEQUENCE OF 1-292 FROM N.A.
RX MEDLINE=95154831; PubMed=7851889;
RA Marchese A., Docherty J.M., Nguyen T., Heiber M., Cheng R.,
RA Heng H.H.O., Tsui L.-C., Shi X., George S.R., O'Dowd B.F.;
RT "Cloning of human genes encoding novel G protein-coupled receptors."
RL Genomics 23:609-618(1994).
CC -1- FUNCTION: ORPHAN RECEPTOR. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE CENTRAL NERVOUS

```



```

RESULT 13
GPCR_MOUSE
ID GPCR_MOUSE STANDARD; PRT; 334 AA.
AC F35412;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE G-PROTEIN-COUPLED RECEPTOR GPCR12 (GPCR01).
DE GPCR12 OR GPCR12
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/C; TISSUE=Brain;
RC MEDLINE=94085630; PubMed=8262253;
RX Saeki Y., Ueno S., Mizuno R., Nishimura T., Fujimura H., Nagai Y.,
RA Wangshihara T.;
RT "Molecular cloning of a novel putative G protein-coupled receptor
RT (GPCR21) which is expressed predominantly in mouse central nervous
RT system."
RL FEBS Lett. 336:317-322(1993).
RL -!- FUNCTION: ORPHAN RECEPTOR.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE FOREBRAIN AND A
CC LESSER EXTENT IN THE HINDRAIN. LOWER EXPRESSION IN THE LIVER.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D21061; BAA04640.1;
CC GCRDb: GCR_0502;
CC MGD: MGI:101509; Gpcr12.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm_1;
CC PRINTS: PR00644; GPCRORPHANR.
CC PRINTS: PR00650; GPCR12ORPHANR.
CC PROSITE: PS00237; G_PROTEIN_RECEPT_FL_1;
CC PROSITE: PS50532; G_PROTEIN_RECEPT_FL_2;
CC -----

```















	Homo sapiens.
XX	Human PFI-006 encoding nucleotide sequence SEQ ID NO:1.
DE	
XX	
KW	Anuretic; PFI-006; G-protein coupled receptor; GPCR; chromosome 19;
KW	anorectic; antidiabetic; neuroprotective; neuroleptic; nootropic;
KW	cystopathic; antiinflammatory; cytostatic; vulnary; dermatological;
KW	osteopathic; cardiovascular; antiallergic; gene therapy; diabetes;
KW	G-protein coupled receptor modulator; signal transduction; obesity;
KW	metabolic disease; neurological disease; psychotherapeutic; inflammation;
KW	urogenital disease; reproduction; sexual medicine; cancer; tissue repair;
KW	dermatology; skin pigmentation; photoaging; frailty; osteoporosis; ds;
KW	cardiovascular disease; gastrointestinal disease; antifunction; allergy;
KW	respiratory disease; sensory organ disorder; sleep disorder; hair loss.
OS	
XX	
FH	Key Location/Qualifiers
CDS	1..1198
FT	/tag= "PFI-006"
FT	/product= "putative G-protein coupled receptor"
FT	/note= "putative G-protein coupled receptor"
XX	
PX	EPI090925-AI.
PD	** 11-APR-2001:
XX	
PF	06-OCT-2000; 2000EP-0308853.
PR	08-OCT-1999; 99GB-0023890.
PA	(PFIZ ) PFIZER LTD.
PA	(PFIZ ) PFIZER INC.
XX	
PI	Harland L;
DR	WI: 2001-302043/32.
DR	P-PADB; AAB74956.
XX	
PT	New human G-protein coupled receptor polynucleotide and polypeptide,
PT	useful in drug screening, as well as for diagnosing or treating
PT	diseases associated with signal transduction, e.g. obesity,
PT	Inflammation, tissue repair or cancer .
PS	Claim 1; Page 36; 46pp; English.
CC	The present sequence encodes human PFI-006 which is a putative G-protein
CC	coupled receptor (GPCR). PFI-006 was identified from chromosome 19.
CC	PFI-006 has anorectic, antidilactic, neuroprotective, neuroleptic,
CC	nootropic, psychotropic, anti-inflammatory cytosolic, vulnarly,
CC	dermatological, osteopathic, cardiovascular and antiallergic activities.
CC	It can be used in pharmaceuticals or in the manufacture of medicaments
CC	for the treatment of a patient having need to modulate, i.e. agonize or
CC	(selectively) antagonize PFI-006 protein. A compound which modulates
CC	PFI-006 can be used in gene therapy. PFI-006 polynucleotides and
CC	proteins are also useful for evaluating and/or screening agents that can
CC	modulate the GPCR. In particular, they are useful for screening drugs
CC	that treat diseases associated with signal transduction, including
CC	obesity, diabetes and metabolic disease, neurological disease,
CC	psychotherapeutics, urogenital disease, reproduction and sexual
CC	medicine, inflammation, cancer, tissue repair, dermatology, skin
CC	pigmentation, photoaging, frailty, osteoporosis, cardiovascular disease,
CC	gastrointestinal disease, antifunction, allergy and respiratory
CC	disease, sensory organ disorders, sleep disorders or hair loss.
XX	
SQ	Sequence 1198 BP; 139 A; 449 C; 389 G; 221 T; 0 other;
	Query Match            99.8%; Score 1195; DB 22; Length 1198;
	Best Local Similarity 100.0%; Pred. No. 5.3e-216;
	Matches 1195; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Oy	i atggagtcgggtctgcgcccgccgcgcgaatcgcacgatcatttgcataaac 60

Db 1081 agcggctcggagcgtctatcgcccccagcgcgagcggtggagaccagcgcggtccacagcc 1140  
 QY 1141 agccccgggtgacccacacaccccccggactctggtatcagaaccggctgcagact 1195  
 Db 1141 agccccgggtgacccacacaccccccggactctggtatcagaaccggctgcagact 1195

RESULT 4  
 AAZ61489  
 ID AAZ61489 standard; DNA; 1980 BP.  
 AC AAZ61489;  
 XX  
 DT 19-JUN-2000 (first entry)  
 DE DNA encoding a 14274 receptor protein.  
 XX  
 KW G-protein coupled receptor; EDG receptor; 14274 receptor;  
 KW spleen disorder; splenomegaly; lung disorder; emphysema;  
 KW chronic bronchitis; colon disorder; atresia; stenosis; enterocolitis;  
 KW liver disorder; jaundice; cholestasis; brain disorder; cerebral edema;  
 KW cerebrovascular disease; Alzheimer's disease; Parkinson's disease;  
 KW T-cell disorder; autoimmune disease; transplant rejection; skin disease;  
 KW vitiligo; lentigo; malignant melanoma; heart disorder; breast disorder;  
 KW ischemic heart disease; kidney disorder; cystic disease; inflammation;  
 KW mastitis; mammary duct ectasia; prostate disorder; benign enlargement;  
 KW tumour; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 152..1348  
 FT CDS /\*tag= .a  
 FT /product= "14274 receptor"  
 PN WO200011166-A1.  
 RD 02-MAR-2000.  
 XX  
 PF 19-AUG-1999; 99WO-US18976.  
 XX  
 PR 19-AUG-1998; 98US-0136726.  
 PR 19-AUG-1999; 99US-0377429.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Glucksmann MA, Welch NS, Hunter JJ;  
 XX  
 DR WPI; 2000-224699/19.  
 DR P-PSDB; AAY69384.  
 XX  
 PT New G-protein coupled receptor used in receptor assays as a target for  
 PT diagnosis and treatment of receptor-mediated disorders of the spleen,  
 PT lung, brain, heart and liver -  
 XX  
 PS Claim 3; Fig 1; 102pp. English.  
 XX  
 CC The present sequence encodes a polypeptide which is a G-protein coupled  
 CC receptor (GPCR) related to the EDG (undefined) receptor family, and is  
 CC designated the 14274 receptor. The receptor polypeptides are useful for  
 CC producing antibodies specific for the GPCR 14274 receptor protein, and  
 CC in drug screening assays, in cell-based or cell-free systems. The  
 CC receptor polypeptide may also be used to identify compounds that  
 CC modulate receptor activity, that is compounds which stimulate or  
 CC inhibit interaction between the receptor protein and a target molecule.  
 CC Modulators of receptor protein activity are used to treat subjects with  
 CC a disorder mediated by the GPCR 14274 receptor pathway, such as;  
 CC as emphysema, chronic bronchitis, disorders of the colon (such as  
 CC atresia, stenosis, enterocolitis), disorders of the liver (such as  
 CC jaundice, cholestasis), disorders of the brain (such as cerebral edema,  
 CC cerebrovascular disease, Alzheimer's disease, Parkinson's disease),  
 CC disorders involving T-cells (such as autoimmune diseases, transplant

CC rejection), skin diseases (such as vitiligo, lentigo, malignant  
 CC melanoma), heart disorders (such as ischemic heart disease), kidney  
 CC disorders (such as cystic diseases), disorders of the breast (such as  
 CC inflammations, mastitis, mammary duct ectasia), and prostate disorders  
 CC (such as inflammations, benign enlargement, tumours).  
 XX  
 SQ Sequence 1980 BP; 340 A; 611 C; 620 G; 392 T; 17 other;

Query Match 99.7%; Score 1193.8; DB 21; Length 1980;  
 Best Local Similarity 99.8%; Pred. No. 9e-216;  
 Matches 1195; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atggagtcggggctgctgcggcgccggcgagcggagtcgcctgcattacaac 60  
 Db 152 atggagtcggggctgctgcggcgccggcgagcggagtcgcctgcattacaac 211  
 QY 61 tacacggcgaagctcgcgggtgcgcgtaccagcgggtgcgcgtccgcgcgcgcgcgc 120  
 Db 212 tacacggcgaagctcgcgggtgcgcgtaccagcgggtgcgcgtccgcgcgcgcgcgc 271  
 QY 121 gtggtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 180  
 Db 272 gtggtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 331  
 QY 181 ctgcgacgcacccgccttcacgcgtccatgcttctgctgctgctgctgctgctgctgct 240  
 Db 332 ctgcgacgcacccgccttcacgcgtccatgcttctgctgctgctgctgctgctgctgct 391  
 QY 241 tcggatctgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 300  
 Db 392 tcggatctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 451  
 QY 301 ctgaaactgtcccccgcgtctgtgtgctgctgctgctgctgctgctgctgctgctgct 360  
 Db 452 ctgaaactgtcccccgcgtctgtgtgctgctgctgctgctgctgctgctgctgctgct 511  
 QY 361 gcgtcgtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 420  
 Db 512 gcgtcgtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 571  
 QY 421 gggccgcgcgcgtctcagtcgggggcgcgcgtgctgctgctgctgctgctgctgctgctg 480  
 Db 572 gggccgcgcgcgtctcagtcgggggcgcgcgtgctgctgctgctgctgctgctgctgctg 631  
 QY 481 gtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 540  
 Db 632 gtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 691  
 QY 541 gcttgctcactgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 600  
 Db 692 gcttgctcactgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 751  
 QY 601 ttcgtgggcacccgcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 660  
 Db 752 ttcgtgggcacccgcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 811  
 QY 661 gccaacgcgcgcgcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720  
 Db 812 gccaacgcgcgcgcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 871  
 QY 721 cgtgcgaagcgcgcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 780  
 Db 872 cgtgcgaagcgcgcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 931  
 QY 781 gtggcatgttggtggccccccttctgctgctgctgctgctgctgctgctgctgctgctgctgctg 840  
 Db 932 gtggcatgttggtggccccccttctgctgctgctgctgctgctgctgctgctgctgctgctgctg 991  
 QY 841 acctgtcctgtactcctgcagggccgagtccttctgctgctgctgctgctgctgctgctgctg 900  
 Db 992 acctgtcctgtactcctgcagggccgagtccttctgctgctgctgctgctgctgctgctgctg 1051



















[illegible]





	QY	277	a t e c t a c t g t c g g g c c g c t c a c g c t g a a a c t g t c c c c g c g c t c t g t t c g c a c g g a g	336
	Dd	304	c t g c t c t g t c g g g c c a c a c t a c a a g c t c a c t c c c g c c a g t g t t c t c g c g g a a	363
	QY	337	g g a g c g c t c t c g t g g c a c t c a c t c g c t c g c t g a c c t c c t g c c a t c g c g c t g g a g	396
	Dd	364	g g a g t a t g t t t g g c c t g t c a g c t c c g t g t c o a g t c t c c t c g c a c t c g c a t t g a g	423
	QY	397	c g a c c c t c a c c a t g g c g c g a g g g g c c c g c c c t c t c a g t c g g g g c g c a c g t g	456
	Dd	424	c g c t a t a c a a t g c t g a a a t g a a a c t c c a a c g g a g a a a a c t t c c g c c t c t c	483
	QY	457	g c g a t g g a c g c g c c g c t g g g c g t g c g t c c t c g c t c c t g g c t c g c a g c g c t g g g c	516
	Dd	484	c t g c a a t a a g c g c t g t g g c t c a t c c c c a c t c c c g g t g g c c t c a t c a t c a t c a g g c	543
	QY	517	t g g a a t t c c t g g g t c c c t g g a c g c t t g c t c a c t g t c t t g c g c t c t a c g c a a a g g c s	576
	Dd	544	t g g a a c t g a t c a g t g c g c t g c c a g c t g c t c a c c g t g c g c g t c a c c a a a g a c	603
	QY	577	t a c g t g c t c t c g t g t c g c t c t g t g g c a t c c t g g c o r e t a t e t g t g a c t c a t a c	636
	Dd	604	t a t a c c c t c t c a c a c g c t c a c t c t g t c t g c t c t c a t c g t c a t t c t g t a c	663
	QY	637	g c g c a t c t a c t c c a g t a c g c g a a c g c g g c c c t g c g g c a c g c c c g g a c t	696
	Dd	664	t g c a g a a t c a c t c c t l g g t c a g a c t c g a g c c c g c c t g a c g t c c g a a g a a c a t t	723
	QY	697	g c g g g a c a c c t g a c c c g g c g c t c a a g c c g c t g c t g y c c t g t g c g c a g	756
	Dd	724	t c c a a g c c a g c-----c g a c t c t g a a g t g c t g g c g t g c t c a a g a c	771
	QY	757	c t c a c g t g g t c c t g c c t t t g g c a t g t g g g c c c c t c t c a t g c t g c t g t g	816
	Dd	772	g t a a t t a c t c c t g a g c g t c t c a t g c t c t g g g a c c g c t c t c a c c t g c t c c g t	831
	QY	817	c t c a g t g g c g t c c c g g c g c a c c t g c t g t a c t c t g a g g c c g a t c c c t c c t g	876
	Dd	832	c t g g a t g t g g g t c a a g g t g a a c c t g a c a c c t c t c a c c t c t c a g a g c g a g t a c t c c g	891
	QY	877	g p a c t g g c a t g g c a a c t c a c t t o t g a a c c c a c a t e t a c a c g c t c a c c a a c c g a c	936
	Dd	892	g t g t a g t g t g c t a a c t c g g c a c a a c c c a c a t t a c a c t c t g a c a a a g g a g	951
	QY	937	c t g c g c a c g c g c t c t g c c t g g t c t g t c g g a c g c c a c t c t g g	986
	Dd	952	a t g c g t g g c c t c a t c c g g a t c a t g t c t c t g c a a g t a c c g a c g	1001

RESULT	15
AAX82789	
ID	AAX82789 standard; DNA; 1155 BP.
XX	
AC	AAX82789;
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	Human edg6 DNA.
XX	
KW	edg6; human; G-coupled receptor; endothelial differentiation gene;
KW	antiflammatory; immunomodulatory; antimicrobial; antiallergic;
KW	Cytostatic; gene therapy; inflammation; autoimmune disease; allergy;
XX	tumor; leukemia; lymphoma; ss.
XX	
OS	Homo sapiens.
XX	
PN	DE19846979-A1.
XX	
PD	23-MAR-2000.
XX	
PF	13-OCT-1998; 98DE-1046979.
XX	
PR	11-SEP-1998; 98DE-1043240.

(DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

Graeler M, Bernhardt G, Lipp M;

WPI: 2000-258069/23.

P-PSDB; AA#90862.

New human and murine G-coupled receptor EDG (endothelial differentiation gene) 6, useful for modulating inflammatory and immune reactions and for treatment of allergy or tumors

Claim 4; Page 7; 12pp; German.

This invention describes novel human and murine G-coupled receptors EDG (endothelial differentiation gene) 6 (I and II). The products of the invention have antiinflammatory, immunomodulatory, antimicrobial, antiallergic and cytostatic activity. (I) and (II) are involved in signal transduction. (I), (II) and their fragments, variants and mutants or binding partners, are used therapeutically to modulate the function of blood and body cells, particularly for inhibition of acute and chronic inflammation and to raise specific antibodies against them. They are used as a source of diagnostic oligonucleotides and for gene therapy.

Antibodies against (I) and (II) are useful for diagnosis and optionally when coupled to therapeutic agents, toxins or other antibodies, to modulate immune and inflammatory responses for example immunological defects such as inflammation, infection, autoimmune diseases, allergy, tumors, leukemia and lymphoma. This sequence encodes the human EDG6 protein described in the method of the invention.

Sequence 1155 BP; 139 A; 414 C; 382 G; 220 T; 0 other:

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein : protein search, using sw model

Run on: January 16, 2002, 22:38:38 ; Search time 73.84 Seconds

(without alignments)  
399.257 Million cell updates/sec

Title: US-09-842-316-2

Perfect score: 2019

Sequence: 1 MESGLLRAPVSEVIVLHYN.....TGSPGAPTAARTLVSEPAAD 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2019	100.0	398	22	Human nGPCR9 #2.
2	2019	100.0	398	22	Amino acid sequenc
3	2019	100.0	399	22	Human PFI-006 prot
4	2018	100.0	398	21	Amino acid sequenc
5	2013	99.7	500	21	Human orphan G pro
6	2013	99.7	500	21	Human G protein co
7	1714.5	82.9	400	22	Rat G protein-coup
8	865.5	42.9	382	20	Human EDG-2 protel
9	865.5	42.9	382	22	LPA receptor-relat
10	863.5	42.8	382	21	Mouse EDG1 polypep
11	853.5	42.3	383	18	p(rat-edg), G-prot

12	853.5	42.3	383	20	AAW87791	Rat-edg, G-protein
13	842	41.7	381	20	AAI33205	Human EDG-1c recep
14	834	41.3	381	22	AAI14337	G-protein coupled
15	834	41.3	381	12	AAI70503	Angiogenesis prote
16	776	38.4	334	15	AAI48752	Human endothelial
17	776	38.4	334	17	AAI02724	Human endothelial
18	774.5	38.4	378	20	AAI05490	Human EDG-3 protel
19	774.5	38.4	378	20	AAI59176	Human EDG3b polyp
20	774.5	38.4	378	22	AAU00303	LPA receptor-relat
21	750.5	37.2	378	22	AAI1751	Murine EDG3 polype
22	733.5	36.3	379	22	AAI73483	Rat G protein-coup
23	732.5	36.3	352	18	AAI58712	Fragment of recept
24	732.5	36.3	352	15	AAI01663	p(H218), G-protein
25	732.5	36.3	352	20	AAI05492	Human EDG-5 protel
26	732.5	36.3	352	20	AAI87790	Human EDG-5 protel
27	732.5	36.3	352	22	AAU00305	LPA receptor-relat
28	728.5	36.1	353	20	AAI49904	Rat H218, G-protein
29	722	35.8	353	21	AAI18889	Human EDG family H
30	721.5	35.7	353	20	AAI28288	Amino acid sequenc
31	721.5	35.7	353	20	AAI28289	Predicted polypept
32	720	35.7	353	20	AAI49905	EDG-4 amino acid s
33	712.5	35.3	352	21	AAI03966	Human EDG family H
34	647	32.0	384	20	AAI06412	Murine EDG5 polype
35	645.5	32.0	384	20	AAI06411	Human EDG-7 recept
36	645.5	32.0	384	21	AAI69500	A human G-protein
37	645.5	32.0	384	21	AAI09862	Human edg6 protel
38	607.5	30.1	386	21	AAI09863	Murine edg6 protel
39	548	27.1	509	19	AAI80955	Amino acid sequenc
40	518.5	25.7	364	19	AAI5451	Rat olfactory bulb
41	518.5	25.7	364	19	AAI59501	Rat olfactory bulb
42	518.5	25.7	364	20	AAI80990	Murine lysophospha
43	518.5	25.7	364	22	AAI37347	Murine LPA recept
44	508.5	25.2	351	20	AAI82659	Human edg-6 protel
45	507.5	25.1	364	18	AAI07619	Human G-protein co

#### ALIGNMENTS

RESULT 1  
AAG80967  
ID AAG80967 standard; Protein; 398 AA.

XX AAG80967;

XX 28-AUG-2001 (first entry)

XX Human nGPCR9 #2.

XX G protein-coupled receptor; nGPCR; seven transmembrane receptor;  
KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
KW cardiovascular disease; proliferative disorder; hormonal disorder;  
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
KW attention deficit-hyperactivity disorder/attention deficit disorder;  
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
KW neuroprotective.

XX Homo sapiens.

XX WO200136473-A2.

XX 25-MAY-2001.

XX 16-NOV-2000; 2000WO-US31581.

XX 16-NOV-1999; 99US-0165838.

XX 17-NOV-1999; 99US-0166071.

XX 19-NOV-1999; 99US-0166678.

XX 28-DEC-1999; 99US-0173396.

XX 22-FEB-2000; 2000US-0184129.

XX 28-FEB-2000; 2000US-0185421.

PR 28-FEB-2000; 2000US-0185554.  
 PR 02-MAR-2000; 2000US-0186530.  
 PR 03-MAR-2000; 2000US-0186811.  
 PR 09-MAR-2000; 2000US-0188114.  
 PR 17-MAR-2000; 2000US-0190310.  
 PR 21-MAR-2000; 2000US-0190800.  
 PR 20-APR-2000; 2000US-0198568.  
 PR 02-MAY-2000; 2000US-0201190.  
 PR 08-MAY-2000; 2000US-0203111.  
 PR 23-MAY-2000; 2000US-0207094.  
 XX (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;  
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
 XX WPI; 2001-389826/41.  
 DR N-PSDB; AAH51007.  
 XX  
 XX New G protein-coupled receptor (GPCR-x) and its encoding  
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
 XX Claim 37; Page 89; 261pp; English.  
 XX  
 CC The present invention relates to novel G protein-coupled receptors  
 CC (GPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as  
 CC seven transmembrane receptors and function in signal transduction. The  
 CC GPCR coding sequences are useful for screening a human to diagnose a  
 CC disorder affecting the brain or a genetic predisposition, specifically  
 CC schizophrenia. GPCRx are useful for identifying compounds useful for  
 CC treating schizophrenia. Detection of GPCR in a sample is useful as a  
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal  
 CC failure, rheumatoid arthritis, CNS disorders, infectious such as HIV-1,  
 CC metabolic and cardiovascular diseases, proliferative disorders and  
 CC hormonal disorders. Modulators of GPCR activity have the utility for  
 CC treating neurological disorders, including schizophrenia, ADHD/ADD  
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),  
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,  
 CC migraine and senile dementia. Additional disorders include inflammatory  
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune  
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory  
 CC diseases e.g. inflammatory bowel disease.  
 XX  
 SQ Sequence 398 AA;  
 Query Match 100.0%; Score 2019; DB 22; Length 398;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-181;  
 Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MESGLRPAPSEVIVLHNYTGKRGARYQFGAGLRADAVVCLAVCAFIENLAVLLV 60  
 DB 1 mesglirpavsevivlhyntgkrgarypgaglradaavvcavclavlenlavllv 60  
 QY 61 LGRHPRFAPMELLGSLTSLDLAGAAYANILLGSLTLKLSPALWFAREGGVFALT 120  
 DB 61 lgrhprfapmfillgsltsldlagaayanillsgptlklspalwfareggvfvalt 120  
 QY 121 ASVLSLALERSLTARRGPAPVSSRGRTLAMAAWVSLGLLGPALGNCLGRLD 180  
 DB 121 asvlsllalalersltarrgpapvssrgrtlamaaawgvslilgllpalgnclgrld 180  
 QY 181 ACSTVPLVAKAVLCVLAFCVILAATCALYARIYQVQVANAARLPARGTAGTSTRA 240  
 DB 181 acstvlplyakavlfvcvllaatcalyariyqvqvanarripargtagtsttra 240  
 QY 241 RRKPRSLALRLTSVVLAFVACWGPFLLELLDVPARTCPVLLQADPFLGLMANSL 300  
 DB 241 rrpksrslalrltsvvlafvacwgpflillldvacpartcpvllqadpflglmanasl 300  
 QY 301 LNPFIYITNDRDLRHALLRVCCGRHSCGRDPGSGQSSASAAEASGGLRCLPPGLDGSF 360

DB 301 lnpfiytitnrdrlrhallrvccgrhscgrdpsgsqsaasagglrrclppgidgsf 360  
 QY 361 SSSRSRSPQRDLTSGTSGPGAPTAARTLVSEPAAD 398  
 DB 361 sssrsrspqrldtsgtsgpgaptaartlvsepaad 398  
 RESULT 2  
 AAB31665  
 ID AAB31655 standard; Protein: 398 AA.  
 XX  
 AC AAB31655;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 DE Amino acid sequence of a human AXOR29 polypeptide.  
 XX  
 KW Human; G-coupled protein receptor; AXOR29; infection; HIV-1; HIV-2; pain;  
 KW cancer; diabetes; obesity; anorexia; bulimia; asthma; hypotension;  
 KW Parkinson's disease; acute heart failure; hypertension; angina pectoris;  
 KW urinary retention; osteoporosis; myocardial infarction; stroke; ulcer;  
 KW allergy; benign prostatic hypertrophy; migraine; vomiting; schizophrenia;  
 KW psychotic disorder; neurological disorder; anxiety; manic depression;  
 KW depression; delirium; dementia; mental retardation; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 XX MO200104119-A2  
 PD 18-JAN-2001  
 XX  
 PF 13-JUL-2000; 2000WO-US19001.  
 PR 13-JUL-1999; 99GB-0016417.  
 PR 08-DEC-1999; 99US-0169573.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Ames RS, Elshourbagy N, Foley JJ, Michalovich D, Sarau HM;  
 PI Smith R, Tsui P, Vawter L, Agarwal P, Lane P;  
 XX WPI; 2001-147176/15.  
 DR N-PSDB; AAF25128.  
 XX  
 PT Novel G-coupled protein receptor, AXOR29 useful for treating diseases  
 PT such as microbial infections, cancers, obesity, asthma, diabetes,  
 PT hypotension, osteoporosis, myocardial infarction and neurological  
 PT disorders  
 XX  
 PS Claim 1; Fig 2; 41pp; English.  
 CC  
 CC The present sequence represents a human G-coupled protein receptor,  
 CC designated AXOR29. AXOR29 polypeptides and polynucleotides are useful  
 CC for treating and diagnosing infections such as bacterial, fungal,  
 CC protozoan and viral infections, particularly infections caused by  
 CC HIV-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia, bulimia,  
 CC asthma, Parkinson's diseases, acute heart failure, hypotension,  
 CC hypertension, urinary retention, osteoporosis, angina pectoris,  
 CC myocardial infarction, stroke, ulcers, allergies, benign prostatic  
 CC hyper trophy, migraine, vomiting, psychotic and neurological disorders  
 CC including anxiety, schizophrenia, manic depression, depression, delirium,  
 CC dementia and severe mental retardation. The polypeptide is also useful  
 CC as a vaccine for inducing immunological response in a mammal.  
 XX  
 SQ Sequence 398 AA;

Query Match 100.0%; Score 2019; DB 22; Length 398;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-181;  
 Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



KW T-cell disorder; autoimmune disease; transplant rejection; skin disease;  
 KW vitiligo; lentigo; malignant melanoma; heart disorder; breast disorder;  
 KW ischemic heart disease; kidney disorder; cystic disease; inflammation;  
 KW mastitis; mammary duct ectasia; prostate disorder; benign enlargement;  
 KW tumour.  
 XX Homo sapiens.

Key Location/Qualifiers  
 PH 1..39 "extracellular domain"  
 FT 40..62  
 FT /note- "transmembrane domain segment"  
 FT 71..95  
 FT /note- "transmembrane domain segment"  
 FT 114..131  
 FT /note- "transmembrane domain segment"  
 FT 132..134  
 FT /note- "GPCR signature"  
 FT 152..173  
 FT /note- "transmembrane domain segment"  
 FT 192..213  
 FT /note- "transmembrane domain segment"  
 FT 253..273  
 FT /note- "transmembrane domain segment"  
 FT 291..308  
 FT /note- "transmembrane domain segment"  
 FT 309..398  
 FT /note- "intracellular domain"

WO200011166-A1.

02-MAR-2000.

19-AUG-1999; 99WO-US18976.

19-AUG-1998; 98US-0136726.

19-AUG-1999; 99US-0377429.

(MILL-) MILLENNIUM PHARM INC.

Glucksmann MA, Welch NS, Hunter JJ;

WPI: 2000-224699/19.

N-PSDB; AA261489.

New G-protein coupled receptor used in receptor assays as a target for  
 diagnosis and treatment of receptor-mediated disorders of the spleen,  
 lung, brain, heart and liver

Claim 1: Fig 1: 102pp; English.

The present sequence represents a polypeptide which is a G-protein  
 coupled receptor (GPCR) related to the EDG (undefined) receptor family,  
 and is designated the 14274 receptor. The receptor polypeptides are  
 useful for producing antibodies specific for the GPCR 14274 receptor  
 protein, and in drug screening assays, in cell-based or cell-free  
 systems. The receptor polypeptide may also be used to identify compounds  
 that modulate receptor activity, that is compounds which stimulate or  
 inhibit interaction between the receptor protein and a target molecule.  
 Modulators of receptor protein activity are used to treat subjects with  
 a disorder mediated by the GPCR 14274 receptor pathway, such as;  
 disorders of the spleen (such as splenomegaly), lung disorders (such  
 as emphysema, chronic bronchitis), disorders of the colon (such as  
 atresia, stenosis, enterocolitis), disorders of the liver (such as  
 jaundice, cholestasis), disorders of the brain (such as cerebral edema,  
 cerebrovascular disease, Alzheimer's disease, Parkinson's disease),  
 disorders involving T-cells (such as autoimmune diseases, transplant  
 rejection), skin diseases (such as vitiligo, lentigo, malignant  
 melanoma), heart disorders (such as ischemic heart disease), kidney  
 disorders (such as cystic diseases), disorders of the breast (such as  
 inflammation, mastitis, mammary duct ectasia), and prostate disorders  
 (such as inflammations, benign enlargement, tumours).

XX	Sequence	398 AA;
XX	Query Match	100.0%; Score 2018; DB 21; Length 398;
XX	Best Local Similarity	99.7%; Pred. No. 1.1e-180;
XX	Matches	397; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 NESGLLRPAVSEVIVLHYNTGKLRGARYQFAGLRADAVVCLAVCAFIIVLENLAVLV 60	
DB	1 mesgllrpapvsevivlhyntgkrgaryqpgaglradvclavcafiivlenlavlv 60	
QY	61 LGRHPRFAPMELLSLTSLDLAGAAYANILLSGPLTLKSLPALWFAREGGVFVALT 120	
DB	61 lgrhprfapmllsgltisdllagaayaanillsgpltlklsplwfareggvfvvalt 120	
QY	121 ASVLSLAIATLERSLTMRGAPVSSRGRITLMAAAAGVSLIIGLLPALGWNCGLRLD 180	
DB	121 asvlsilalatersltmarrgppvssrgrtlaamaaagvslilgllpalgwnclgrld 180	
QY	181 ACSTVLPYAKAYVLCVLAFAVGIILAAICALYARIYCOVRANARLPPARCTAGTTSTRA 240	
DB	181 acstvlpyakayvfcvlfafvgiilaicalyariycqiranarrlparpgtagttstra 240	
QY	241 RRKPRSLALRTLSTVLLAFVACWGPLFLLLLLDVACPARTCPVLQADPFLGLAMANS 300	
DB	241 rrkprslalrtlstsvllafvacwgpflflllldvacpartcpvllqadpflglamansl 300	
QY	301 LNPIIYTLNRDLRHALLRLVCCGRHSCGRDPGSGQSAASAAESGGLRRLCPPLDGSF 360	
DB	301 lnpiliytlndrlrhallrlvccgrhscgrdpdsgsgqsasaaesgglrrclppldgsf 360	
QY	361 SGRSSSRQDGLDTSGSTGSPGAPTARTLVSEPAAD 398	
DB	361 sgrsssrqrdgldtsgatgspgaptartlvsepaad 398	
RESULT	5	
XX	AAV71305	
XX	ID AAV71305 standard; Protein; 500 AA.	
XX	AC AAV71305;	
XX	XX AAV71305;	
DT	02-NOV-2000 (first entry)	
DE	Human orphan G protein-coupled receptor hCHN6.	
XX	Human; orphan G protein-coupled receptor; GPCR; hCHN6; drug screening;	
KW	transmembrane receptor; expressed sequence tag; EST; signal cascade.	
XX	XX Homo sapiens.	
XX	WO20001258-A2.	
XX	02-JUN-2000.	
PF	13-OCT-1999; 99WO-US23687.	
XX	20-NOV-1998; 98US-0109213.	
PR	16-FEB-1999; 99US-0120416.	
PR	26-FEB-1999; 99US-0121852.	
PR	12-MAR-1999; 99US-0123946.	
PR	28-MAY-1999; 99US-0123949.	
PR	28-MAY-1999; 99US-0136436.	
PR	28-MAY-1999; 99US-0136437.	
PR	28-MAY-1999; 99US-0136439.	
PR	28-MAY-1999; 99US-0136567.	
PR	28-MAY-1999; 99US-0137127.	
PR	28-MAY-1999; 99US-0137131.	
PR	29-JUN-1999; 99US-0141448.	
PR	29-SEP-1999; 99US-0156555.	
PR	29-SEP-1999; 99US-0156633.	
PR	29-SEP-1999; 99US-0156634.	

PR 29-SEP-1999; 99US-0156653.  
 PR 01-OCT-1999; 99US-0157280.  
 PR 01-OCT-1999; 99US-0157281.  
 PR 01-OCT-1999; 99US-0157282.  
 PR 01-OCT-1999; 99US-0157283.  
 PR 01-OCT-1999; 99US-0157293.  
 PR 01-OCT-1999; 99US-0157294.  
 PR 12-OCT-1999; 99US-0416760.  
 PR 12-OCT-1999; 99US-0417044.  
 PA (AREN-) ARENA PHARM INC.  
 XX Chen R, Dang HT, Llaw CW, Lin I;  
 XX WPI: 2000-400068/34.  
 DR N-PSDB; AAD01132.  
 XX Novel human orphan G protein-coupled receptors and the encoding cDNAs  
 PT for use in the identification of G protein-coupled receptor agonists -  
 PS Claim 58; Page 79-81; 102pp; English.  
 XX The present amino acid sequence is the hCHN6, an endogenous human  
 CC orphan G protein-coupled receptor (GPCR). The hCHN6 cDNA was identified  
 CC using ESTs (expressed sequence tag) AA804531 and 2134670 as a probe.  
 CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane  
 CC alpha helices with an extracellular N-terminus and an intracellular  
 CC C-terminus. However, no endogenous ligands has yet been identified for  
 CC the proteins of the invention. The orphan GPCRs may be used in the  
 CC identification of their endogenous ligands, and to screen potential GPCR  
 CC agonists and antagonists for use as pharmaceutical agents. The proteins  
 CC may also be used in the study of GPCR-mediated signaling cascades, and  
 CC to elucidate their precise role in normal and diseased human conditions.  
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue  
 CC localisation expression analysis to provide information about their  
 CC function in healthy and pathological states.  
 XX Sequence 500 AA:  
 SQ  
 Query Match 99.7%; Score 2013; DB 21; Length 500;  
 Best Local Similarity 99.7%; Pred. No. 4.1e-180;  
 Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MESGLRPAVSEVILVHYNTGKIRGARYQPGAGLRADAVVCLAVCAFIENLAVLLV 60  
 Db 103 mesgllrpapvseivilvhyntgkrgasyqpgaglradaavclavcafienvlavllv 162  
 QY 61 LGRHPRFAPMFLLIGSLTSLDLAGAAYANILLGPTLKLSPALWFAREGGVFAVLT 120  
 Db 163 lgrhprfhapmflligsltdlagaayaanillsgptlklspalwfareggvfvavlt 222  
 QY 121 ASVLSLATALERSLTMRGAPVSSRGRTLMAAAANGVSLLLGLLPALGWNCLGRLD 180  
 Db 223 asvlsllatalersltmrargpavssrgtrtlmaaaaangvslllgllpalgwnclgrld 282  
 QY 181 ACSTVPLVYKAYVLCVLAFLVGLIAALCALYARYICOVNARRLPARPCTAGTTSTRA 240  
 Db 283 acstvlplykayvlfvcvllaflvgliaalcayaryicovnarrrlparpctagttstra 342  
 QY 241 RRPRLSLRLTSLVLLAFVACWGPLELLLLLDVACPARTCPVLLQADPFLGLAMANS 300  
 Db 343 rrprrslrltsslvllafvawcwgplellllldvacpartcpvllqadpflglamansl 402  
 QY 301 LNPPIITLNRDLRALRLVCCGRHSCGRPSGSGSOASAAEASGGLRCLPGLDGSF 360  
 Db 403 lnpplittlndlrallrvccgrhscgrpsgsgsoasaaesgglrrcplpdlgsf 462  
 QY 361 SGRSSPQRDLGDTSGTSGSPGAPTAARTLVSEPAAD 398  
 Db 463 sgrsspqrldgdtsgtsgspgaptartlvsepaad 500

RESULT 6

AAB02839  
 ID AAB02839 standard; Protein; 500 AA.  
 XX  
 AC AAB02839;  
 XX  
 DT 22-AUG-2000 (first entry)  
 XX  
 DE Human G protein coupled receptor hCHN6 protein SEQ ID NO:32.  
 XX  
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;  
 KW identification; agonist; screening; therapeutic; pharmaceutical;  
 KW mutant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000022131-A2  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 13-OCT-1999; 99WO-US24065.  
 XX  
 PR 13-OCT-1998; 98US-0170496.  
 PR 12-NOV-1998; 98US-0108029.  
 PR 27-NOV-1998; 98US-0109213.  
 PR 16-FEB-1999; 98US-0110060.  
 PR 26-FEB-1999; 98US-0120416.  
 PR 12-MAR-1999; 98US-0121852.  
 PR 12-MAR-1999; 98US-0123944.  
 PR 12-MAR-1999; 98US-0123945.  
 PR 12-MAR-1999; 98US-0123946.  
 PR 12-MAR-1999; 98US-0123948.  
 PR 12-MAR-1999; 98US-0123949.  
 PR 12-MAR-1999; 98US-0123951.  
 PR 28-MAY-1999; 98US-0136436.  
 PR 28-MAY-1999; 98US-0136437.  
 PR 28-MAY-1999; 98US-0136439.  
 PR 28-MAY-1999; 98US-0137127.  
 PR 28-MAY-1999; 98US-0137131.  
 PR 28-MAY-1999; 98US-0137567.  
 PR 30-JUN-1999; 98US-0141448.  
 PR 27-AUG-1999; 98US-0151114.  
 PR 03-SEP-1999; 98US-0152524.  
 PR 29-SEP-1999; 98US-0156633.  
 PR 29-SEP-1999; 98US-0156555.  
 PR 29-SEP-1999; 98US-0156634.  
 XX  
 PA (AREN-) ARENA PHARM INC.  
 XX  
 PI Behan DP, Lehmann-Brulnsma K, Chalmers DT, Chen R, Dang HT;  
 PI Gore M, Llaw CW, Lin I, Lowitz K, White C;  
 XX  
 WPI: 2000-317986/27.  
 DR N-PSDB; AAA46033.  
 XX  
 PT Non-endogenous, human G protein-coupled receptors for screening  
 PT receptor, inverse or partial agonists useful as therapeutic agents  
 XX  
 PS Example 1; Page 109-111; 187pp; English.  
 XX  
 CC The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCR), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,  
 CC inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 500 AA;  
 Query Match 99.7%; Score 2013; DB 21; Length 500;

Best Local Similarity 99.7%; Pred. No. 4,1e-180;  
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESGLLRPAPVSEVILHNYTKLRGARYQPGAGLRADAVVCLAVCAFIIVLENLAVLLV 60  
DB 103 mesgllrpapvsevivlhyntgklrgayqpgaglradaavclavcafiivlenlavllv 162

QY 61 LGRHPRFAPMFLLSGLTSLDLAGAAYAAANLLSGPLTKLSPALWFAREGGVFVALT 120  
DB 163 lgrhprfapmflilsgltslldagaayaanllsgpltklspalwfareggvfvalt 222

QY 121 ASVLSLLAIALERSLTMRGPPAPVSSRGRTLAMAAAGVSLLLGLLPALGWNCLGRLD 180  
DB 223 asvlsllaialersltmrpgpavssrgtrtlaaaagvslldglpallgwnclgrld 282

QY 181 ACSTVPLXYAKAVLFCVLAFAVIGLAAICALLYARIYCOVRANARLPARGTAGTTSTRA 240  
DB 283 acstvpilyakavilfcvlfavlglaaicalyariycovranarrlpargtagttstra 342

QY 241 RRPKPSLALLRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANS 300  
DB 343 rrpkslallrtlsvllafvacwgpflfllllldvacpartcpvllqadpflglamansl 402

QY 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPSSQSQSASAAEASGGRLRCLPPLDGSF 360  
DB 403 lnpiiytltnrdlrhallrvccgrhscgrdpssgsgsasaaesggllrclppldgsf 462

QY 361 SGSESPQRDLDTSGTSGSPCAPTAARTLVSEPAAD 398  
DB 463 sgseस्पqrldtsgtsgspcaptaartlvsepaad 500

RESULT 7  
AAV72561  
ID AAY72561 standard; Protein; 400 AA.  
AC AAY72561;  
XX  
DT 02-MAY-2001 (first entry)  
XX  
DE Rat G protein-coupled receptor (GPCR), B1C3.  
XX  
KW Rat; G protein-coupled receptor; GPCR; B1C3; intracellular signal;  
XX  
OS analgesic; anaesthetic; drug screening.  
XX  
Rattus sp.  
XX  
WO200105829-A1.  
XX  
PD 25-JAN-2001.  
XX  
PF 20-JUL-2000; 2000WO-SE01505.  
XX  
PR 21-JUL-1999; 99SE-0002763.  
XX  
PA (ASTR ) ASTRAZENECA AB.  
XX  
PI Ahmad S, Hoffert C, O'donnell D, Pelletier M, Walker P;  
XX  
WPI; 2001-159515/16.  
DR  
N-PSDB; AAD02561.  
XX  
PT New G protein-coupled receptor, B1C3, expressed in the central nervous  
PT system of rats, for screening modulators of the receptor and  
PT identifying agents of therapeutic use capable of binding to the  
PT receptor -  
XX  
XX Claim 1; Fig 2; 35pp; English.  
PS  
PS The present sequence is rat G protein-coupled receptor (GPCR), B1C3  
CC which is expressed in the central nervous system of rats.  
CC B1C3 nucleic acids and recombinant proteins are useful in assays to  
CC identify agents capable of binding to the receptor, which modulate

CC intracellular signalling such as adeny cyclase activity or  
CC intracellular calcium concentration. These agents have potential  
CC therapeutic applications as either analgesics or anaesthetics. B1C3  
CC receptors are also useful for screening drug candidates using cell  
CC signalling assays. Antibodies to B1C3 are useful in purification of  
CC intact receptor or fragments of the receptor.  
XX  
SQ Sequence 400 AA;

Query Match 84.9%; Score 1714.5; DB 22; Length 400;  
Best Local Similarity 87.2%; Pred. No. 2.7e-152;  
Matches 346; Conservative 14; Mismatches 34; Indels 3; Gaps 3;

QY 1 MESGLLRPAPVSEVILHNYTKLRGARYQPGAGLRADAVVCLAVCAFIIVLENLAVLLV 60  
DB 1 mesgllrpapvsevivlhyntgklrgaryqpgaglradaavclavcafiivlenlavllv 60

QY 61 LGRHPRFAPMFLLSGLTSLDLAGAAYAAANLLSGPLTKLSPALWFAREGGVFVALT 120  
DB 61 lgrhprfapmflilsgltslldagaayaanllsgpltklspalwfareggvfvla 120

QY 121 ASVLSLLAIALERSLTMRGPPAPVSSRGRTLAMAAAGVSLLLGLLPALGWNCLGRLD 180  
DB 121 asvlsllaialersltmrpgpavssrgtrtlaaaagvslldglpallgwnclgrld 180

QY 181 ACSTVPLXYAKAVLFCVLAFAVIGLAAICALLYARIYCOVRANARLPARGTAGTTSTR 239  
DB 181 acstvpilyakavilfcvlfavlglaaicalyariycovranarrlpargtagttssr 240

QY 240 AERKPSLALLRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANS 299  
DB 241 szhrprsalrtlsvllafvacwgpflfllllldvacpartcpvllqadpflglamans 300

QY 300 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPSSG-QQSASAAEAS-GGLRRLCLPPGLD 357  
DB 301 lnpiiytltnrdlrhallrvccgrhscgrdpssnslqrpasvpgsgglrrclppltd 360

QY 358 GSFSGSERSSQQRDLDTSGTSGSPCAPTAARTLVSE 394  
DB 361 rsspsehscqqrldtsgtsgspcaptaartlvse 397

RESULT 8  
AAV05489  
ID AAY05489 standard; Protein; 382 AA.  
XX  
AC AAY05489;  
XX  
DT 07-JUL-1999 (first entry)  
XX  
DE Human EDG-2 protein sequence.  
XX  
KW EDG-1; EDG-2; EDG-3; EDG-4; EDG-5; PSP-24; human; detection; therapy;  
KW inverse agonist; allosteric modulator; lysophosphatidic acid receptor;  
KW LPA signalling mediated disease; cellular apoptosis.  
XX  
OS Homo sapiens.  
XX  
PN WO9919513-A2.  
XX  
PD 22-APR-1999.  
XX  
PF 09-OCT-1998; 98WO-US21315.  
XX  
PR 10-OCT-1997; 97US-0061572.  
XX  
FA (LXRB-) LXR BIOTECHNOLOGY INC.  
XX  
XX Erikson J, Goddard JG, Kiefer M;  
XX WPI; 1999-277658/23.  
DR  
N-PSDB; AAX36566.





Dt	06-FEB-2001 (first entry)
Xx	Mouse EDG1 polypeptide.
De	Mouse; EDG1; antimicrobial; cytostatic; analgesics; antidiabetic;
Xx	antiparkinsonian; cardiant; osteopathic; anorectic; antiasthmatic;
Kw	antimetabolic; tranquiliser; antiallergic; neuroleptic; antitumor;
Kw	antidepressant; infection; cancer; diabetes; eating disorder;
Kw	asthma; Parkinson's disease; heart failure; hypotension; hypertension;
Kw	osteoporosis; angina pectoris; stroke; ulcer; allergy;
Kw	benign prostatic hypertrophy; migraine; vomiting; neurological disorder.
Xx	Mus musculus.
Os	
Pn	WO200059529-A1.
Pd	12-OCT-2000.
Pf	05-APR-2000; 2000WO-US09075.
Pr	05-APR-1999; 99US-0127696.
Pr	04-APR-2000; 2000US-0127896.
Xx	(SMIK ) SMITHKLINE BEECHAM CORP.
Fa	Tau1 P;
Pi	WPI; 2000-664960/64.
N-PSDB;	AAC63385.
Dr	Murine EDG1 polypeptide useful for treating infections and diseases
Xx	such as cancer, diabetes, Parkinson's disease, allergies, and
Pt	neurological disorders such as anxiety schizophrenia, depression -
Xx	Claim 1; Page 31-32; 34pp; English.
Xx	The present sequence is mouse EDG1. Agonists and antagonists of
Cc	mouse EDG1 are useful for treating patients with enhanced
Cc	or inhibited activity or expression of EDG1. The EDG1 polypeptide is
Cc	useful for treating abnormal conditions such as infections caused by
Cc	bacteria, fungi etc., pain, cancer, diabetes, obesity, anorexia, bulimia,
Cc	asthma, Parkinson's disease, acute heart failure, hypotension,
Cc	hypertension, urinary retention, osteoporosis, angina pectoris,
Cc	myocardial infarction, stroke, ulcers, allergies, benign prostatic
Cc	hypertrophy, migraine and vomiting. It is useful for treating psychotic
Cc	and neurological disorders including anxiety, schizophrenia, manic
Cc	depression, delirium, dementia and severe mental retardation
Cc	and dyskinesias such as Huntington's disease or Gilles de la Tourette's
Cc	syndrome.
Xx	Sequence 382 AA;
Sq	
Query Match	42.8%; Score 863.5; DB 21; Length 382;
Best Local Similarity	48.5%; Pred. No. 1.2e-72;
Matches 181;	Conservative 65; Mismatches 106; Indels 21; Gaps 6;
Qy	13 EVIVLHNTYTKGLR-CARVOPGAGLRADAVVCCLAVCAFIENLAVLVGRHRPFHPM 71
Db	23 diivrhnytgklniga--ekdhgikltsvvfilccfillenifvlitwktkkfrpm 80
Qy	72 FLLGSLSLSDLAGAAYANILLSPGLTKLPALWFAREGGVFVALTSVLSLAIAL 131
Db	81 yyfignlalsdlagvaytanillsgattkyktlpaqwflregsmfvaisavfllaia 140
Qy	132 ERSLTWARGPAPVSRRGTFLMAAAANGCVSLLGLLPALGNCLGRDACSTVLPLYAK 191
Db	141 erytlmknklhngsnssrfflliscawvisilggplmgwncisssscstvlplyhk 200
Qy	192 AYVLCVAFVGILAAICARYIQCVRANARRLPARGTAGTSTRARRKPRLALLR 251
Db	201 hylftctvtilllsivlyoriyalvrtstrritfk-----nlksasrsesksalalk 256

Qy	252 TISVVLLAFVACWGPFLLILLIDVACPARTCPVLILQADPPFLGLMANSLNPIIYTITNR 311
Db	257 tviilvsfiacwaplfllilldvckaktcdilykaeyflavlansgtnpilittlk 316
Qy	312 DLRHALLRLVCCGRHSRCRDPSSGOOSASAEAGSLRCPLPGLDGSGFSGSRSS-PQR 370
Db	317 emtrafiriv-----sckcpng-----dsagkkrpilpgmfefarsksdnshpqk 363
Qy	371 DCLDTSGTGSPG 383
Db	364 dëgdnpetimssg 376
RESULT 11	
AAW01664	
ID	AAW01664 standard; Protein; 383 AA.
XX	
AC	AAW01664;
XX	
DT	01-APR-1997 (first entry)
DE	p(rat-edg), G-protein coupled receptor.
XX	
KW	p(H218); G-protein coupled receptor; cell differentiation; proliferation;
KW	proline directed kinase; cell division; growth factor response; rat-edg;
KW	therapy; diagnosis.
XX	Rattus rattus.
OS	
XX	
PN	US5585476-A.
XX	
PD	17-DEC-1996.
PF	15-FEB-1994; 94US-0196989.
XX	
PR	15-FEB-1994; 94US-0196989.
XX	
PA	(MACL/) MACLENNAN A J.
XX	
PI	MacLennan AJ;
XX	
DR	WPI; 1997-051235/05.
DR	N-PSDB; AAT58506.
XX	
PT	DNA encoding rat protein p(H218) - associated with cell
PT	proliferation and/or differentiation
XX	
PS	Example 9; Column 17-20; 33pp; English.
XX	
CC	This sequence is a novel rat protein p(rat-edg), a member of the
CC	G-protein coupled receptor (GPR) superfamily of proteins. The
CC	amino acid similarity between p(H218) (AAW01663) and p(rat-edg) suggests
CC	that they may be activated by the same endogenous ligand(s). The
CC	expression pattern of mRNA transcripts of both genes in cell lines,
CC	various rat tissues and developing rat brain suggests that they both
CC	play a role in cell proliferation and/or differentiation.
XX	
SQ	Sequence 383 AA;
Query Match	42.3%; Score 853.5; DB 18; Length 383;
Best Local Similarity	47.6%; Pred. No. 1e-71;
Matches 177;	Conservative 66; Mismatches 110; Indels 19; Gaps
Qy	13 EVIVLHNTYTKGLRGARYQPAGLRADAVVCCLAVCAFIENLAVLVGRHRPFHPMF 72
Db	24 diivrhnytgkl-nigvekdghiktstvfilccfillenifvlitwktkkfrpm 82
Qy	73 LUGSLTLSDLAGAAYANILLSPGLTKLPALWFAREGGVFVALTSVLSLAIALE 132
Db	83 yfgnlalsdlagvaytanillsgattkyktlpaqwflregsmfvaisavfllaiale 142
Qy	133 RSLTMRRGPAPVSSRRGTFLMAAAANGVSLLGLLPALGNCLGRDACSTVLPLYAKA 192

Db 143 ryltmkmlhngsnssr:fliliscawisililgplimgwncisslscstvlplyhkh 202  
QY 193 YVLCVLAFAVGILAAICALYARIYCOVRANARLPARPCTAGTTSTRARRKPRSLALLRT 252  
Db 203 yilfctvtllililsvilycrylsvtrrrlfrk-----niskasrseksalallkt 258  
QY 253 LSVLLAFVACWGLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRD 312  
Db 259 viilsvfiacwapiifillldvgckaktcdilykaeyflvavlnsgtnpilytltnke 318  
QY 313 LRHALLRLVCCGRHSCGRDPGSGSQSASAAEASGLRLRCLPPGLDGSFSGSERSS-PORD 371  
Db 319 mrrafirli-----sckcpng-----dsagkfrpiipgmefsrsksdnshpqkd 365  
QY 372 GLDTSSTGSPG 383  
Db 366 dgdnpetlmssg 377  
RESULT 12  
AAW87791  
ID AAW87791 standard; Protein; 383 AA.  
XX  
AC AAW87791;  
XX  
DT 12-MAR-1999 (first entry)  
XX  
DE Rat-edg, G-protein coupled receptor superfamily member.  
XX  
KW H218; G-protein coupled receptor superfamily; cell proliferation;  
KW cell differentiation; cancer; rat-edg; diagnosis; therapy.  
XX  
OS Rattus sp.  
XX  
PN US5856443-A.  
XX  
PD 05-JAN-1999.  
XX  
PF 06-DEC-1996; 96US-0760936.  
XX  
PR 15-FEB-1994; 94US-0196989.  
XX  
PR 06-DEC-1996; 96US-0760936.  
XX  
PA (MACL/) MACLENNAN A J.  
XX  
PI MacLennan AJ;  
XX  
DR WPI; 1999-105192/09.  
XX  
DR N-PSDB; AAV84036.  
XX  
PT New isolated polynucleotide encoding a G-protein coupled receptor -  
PT used to develop products for use in diagnosis and therapy of  
PT conditions involving cell proliferation and differentiation  
XX  
PS Disclosure; Columns 23-26; 33pp; English.  
XX  
CC The present sequence represents a protein designated rat-edg. The rat-edg  
CC polypeptide is a member of the G-protein coupled receptor superfamily.  
CC The specification also describes a protein designated H218, which  
CC is also a member of the G-protein coupled receptor superfamily. H218  
CC is involved in cell proliferation and differentiation, and in disease  
CC states such as cancer. The polynucleotides and polypeptides can be  
CC used to develop products for diagnosis and therapy.  
XX  
SQ Sequence 383 AA;  
Query Match 42.3%; Score 853.5; DB 20; Length 383;  
Best Local Similarity 47.8%; Pred. No. 1e-71;  
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;  
QY 13 EVIVLHNYTGKLRGARYQPGAGLRADAVVCLAVCAFIIVLENLVLVLRPHRFAPMF 72

Db 24 diivrhnytkl-nlgvekdghgikltsvffilicollifenfvlttwktkkfrhpy 82  
QY 73 LLGSLTSLDLAGAAYANALLSGPLTKLSPALWFAREGGVFVALTASVLSLLAIALE 132  
Db 83 yfignalaledilagvaytanillsgattykltpeaqwflregsmfvalsaavfallaiaie 142  
QY 133 RSLTWARGPAPVSSRGFTLAMAAANGVSLLLGLLPALGNCLGRDADACSTVPLIYAKA 192  
Db 143 ryltmkmlhngsnssr:fliliscawisililgplimgwncisslscstvlplyhkh 202  
QY 193 YVLCVLAFAVGILAAICALYARIYCOVRANARLPARPCTAGTTSTRARRKPRSLALLRT 252  
Db 203 yilfctvtllililsvilycrylsvtrrrlfrk-----niskasrseksalallkt 258  
QY 253 LSVLLAFVACWGLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRD 312  
Db 259 viilsvfiacwapiifillldvgckaktcdilykaeyflvavlnsgtnpilytltnke 318  
QY 313 LRHALLRLVCCGRHSCGRDPGSGSQSASAAEASGLRLRCLPPGLDGSFSGSERSS-PORD 371  
Db 319 mrrafirli-----sckcpng-----dsagkfrpiipgmefsrsksdnshpqkd 365  
QY 372 GLDTSSTGSPG 383  
Db 366 dgdnpetlmssg 377  
RESULT 13  
AAW87791  
ID AAW87791 standard; Protein; 381 AA.  
XX  
AC AAW87791;  
XX  
DT 18-NOV-1999 (first entry)  
XX  
DE Human EDG-1c receptor protein.  
XX  
KW EDG-1c receptor; sphingosine-1-phosphate; S-1-P; human; anti-arrhythmic;  
KW dihydro sphingosine-1-phosphate; anti-HIV; anti-asthmatic; cytostatic;  
KW anti-angina; analgesic; antimicrobial; neuroprotective; nootropic; pain;  
KW treatment; infection; bacterial; fungal; protozoan; viral; HIV-1; HIV-2;  
KW cancer; diabetes; obesity; anorexia; bulimia; Parkinson's disease;  
KW asthma; acute heart failure; hypertension; hypotension; osteoporosis;  
KW urinary retention; angina pectoris; myocardial infarction; stroke;  
KW congestive heart failure; left ventricular hypertrophy; arrhythmia;  
KW restenosis; coronary artery angioplasty; vascular sclerosis; allergy;  
KW inflammation; deleterious fibrosis; arteriosclerosis; angiogenesis;  
KW wound healing; ulcer; benign prostatic hypertrophy; migraine; vomiting;  
KW psychotic disorder; neurological disorder; anxiety; schizophrenia;  
KW manic depression; depression; delirium; dementia; degenerative disease;  
KW severe mental retardation; neurodegenerative disease; ischemic stroke;  
KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.  
XX  
OS Homo sapiens.  
XX  
PN WO9946277-A1.  
XX  
PD 16-SEP-1999.  
XX  
PF 04-MAR-1999; 99WO-US04675.  
XX  
PR 09-MAR-1998; 98US-0077369.  
XX  
PR 28-MAY-1998; 98US-0087102.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SB LAB PHARM.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX (STAD/) STADEL P.  
PI Bergsma DJ, Chambers JK, Chan W, Johnson RK, Khandoudi N, Livi GP;  
PI Robert P, Stadel JM, Wilson S;





